IBA

TECHNICAL REVIEW

16

Digital Coding Standards

Blank pages in this document were not scanned so there may be occasional gaps in the page sequence.

IBA INDEPENDENT BROADCASTING AUTHORITY

IBA TECHNICAL REVIEW March 1982
ISSN 0308–423 X

16 Digital Coding Standards

Contents

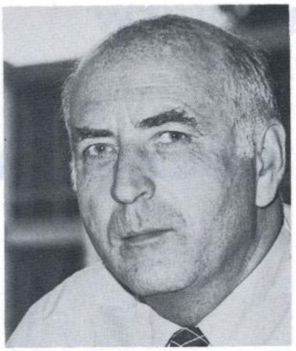
Technical Editor: John Lovell, IBA Engineering Information Service

Additional Copies
Subject to availability, further copies of this *IBA Technical Review* may be obtained on application to Engineering Information Service, IBA, Crawley Court, WINCHESTER, Hampshire, SO21 2QA.
No charge will be made for small quantities.

IBA INDEPENDENT BROADCASTING AUTHORITY

HEADQUARTERS: 70 Brompton Road, LONDON SW3 1EY. *Tel: 01-584 7011; Telex: 24345*
ENGINEERING DIVISION: Crawley Court, WINCHESTER, Hampshire, SO21 2QA. *Tel: 0962 823434; Telex: 477211*

Introduction



by T. S. Robson, OBE
Director of Engineering
Independent Broadcasting Authority

The different television standards that are currently in use today in various parts of the world were all chosen before the advent of the videotape recorder and the dawn of the communications satellite; indeed it is 45 years since the start in the UK of the first public high definition service, as it was then called. This 405-line service is now scheduled to be closed down by 1986 when it will have completed a half century.

If television had not existed prior to the start of communication satellites perhaps we might have had the wisdom to adopt a single world-wide television standard. This, however, is idle speculation; there are two major line standards, three colour systems and a vast investment in receiving equipment. It would be foolhardy to consider changes to the standards radiated by terrestrial transmitters that were not compatible with existing receivers. Indeed, there was warning from over three centuries ago, from the introduction to the 1662 Book of Common Prayer '. . . common experience showeth, that where a change is made of things advisedly established (no *evident necessity* so requiring) sundry inconveniences have thereupon ensued; and those many times more and greater than the evils, that were intended to be remedied by such change'.

Digital processing of the video signal, often in component form, is progressively increasing in the studio and may well become dominant. It has become *evident* that a single digital component standard is a *necessity* to avoid interfacing problems between digital equipment.

Our experience within the IBA of processing digital video signals in component form started in 1971 with the development of the first digital field standards converter (DICE). More recently our work on components has been aimed at contributions towards a world-wide standard for the sampling of component signals within each line of the television picture. In the forthcoming months we are hopeful that concensus on a world-wide standard will increasingly become apparent.

Some of this work towards a single world-wide standard and future possibilities that might ensue are described in this edition of the *IBA Technical Review*; sometimes we invite articles from external authors, this time we are pleased to include one from Mike Tooms.

Tom Robson.

KEITH LUCAS, M.Sc., Ph.D., graduated in 1967, and afterwards spent four years at Southampton University on research in the field of Artificial Intelligence and Adaptive Control Systems. He was then employed by the Plessey Company and worked on a number of defence projects. In 1974 he joined the Authority's Experimental and Development Department where, within the Automation and Control Section, he was engaged in the development of ORACLE. He became Head of Video and Colour Section in the early part of 1978.

The Evolution Towards Component Coded Video Systems

by K. Lucas

Synopsis

In recent months, major steps have been taken towards establishing world-wide agreement on a digital standard for television studio signals. To achieve agreement, and for sound technical and economic reasons, the new standard will be based on the separate component signals which form the basis of the various composite signals in use today. This paper summarises the recent history of the discussions on digital standardisation, and considers the implications of the move towards component coded signals. It concludes that we may expect to see, in future systems, a progressive change from composite to component coding for the distribution, and even for the transmission, of television signals.

Introduction

In 1974, an EBU study group (C1) was established to examine the question of standardisation in the emerging field of digital video. Since then, progress towards an internationally acceptable digital video standard has been erratic, but continuous. The concerted efforts of the EBU and SMPTE are now very close to a successful conclusion; a result which, a few years ago, seemed impossible.

In the early days of colour television, studio systems were fairly simple, and production involved minimal signal processing. The feasibility and the demand for complex production techniques (picture manipulation and special-effects) had been impossible to foresee, and cannot be achieved satisfactorily by use of composite signals (PAL, SECAM or NTSC). Most of the engineers involved in the early discussions of digital video coding could see the long-term advantages of a system based on separate component signals. Such a system would provide improved flexibility and quality in signal processing—particularly in the post-production area. The digital technique essentially removes the need of composite signals up-stream of the transmitting station.

SECAM users have already faced some of the problems presented by the demand for more sophisticated signal processing. The SECAM signal is unsuited for these applications; and separate component signals are already used quite extensively in SECAM studios. The short-term difficulties in adopting a digital code based on component signals would not be very great, neither was there any attractive alternative.

PAL users were in a different situation. The PAL signal provides just sufficient flexibility in signal processing to enable mixing and certain special-effects without decoding to components. The PAL signal is used almost exclusively from camera to transmitter; and equipment based on component signals is largely incompatible, due to the impairments of PAL

decoding. Consequently, a significant short-term advantage would be available to PAL users by adopting a digital code based on composite PAL. Indeed, in the mid-1970s, the adopting of a digital component code was generally considered unfeasible. The digital standards then under study within the EBU are shown in Table 1.

Although the proposed digital PAL and YUV signals seemed to be compatible, this was not entirely the case. The full potential of component coding can be achieved only by adopting static sampling patterns. This provides simplicity in signal processing, and in editing to one-frame resolution. Full compatibility with a sub-carrier-locked sampled PAL signal would demand mobile sample structures (appropriate to the four-frame sub-carrier cycle). It seemed, therefore, that 625-line countries in Europe were favouring the adoption of two digital standards which were compatible only in appearance.

Progress Towards a Common View

The proposals outlined above caused many misgivings among PAL users, for the following main reasons:

(*a*) It was clear that considerations of compatibility in PAL countries were adversely affecting the debate on optimum digital formats. There was a danger that the full advantages of the digital technique would be compromised by short-term considerations.

(*b*) The proposals did not provide a good basis for the development of common 625-line equipment, neither had they any similarities with equivalent proposals in the 525-line domain. Consider, for example, the development of a digital VTR. The 2 fsc PAL proposal required only 71 Mbit/s of data-rate whereas the component signal required 142 Mbit/s. It is true that the recording of 4 fsc PAL would equalise the data-rate for all 625-line users; but advantages which may be gained with digital PAL recording (namely, picture quality) could be achieved at a sample rate of 2 fsc. The 4 fsc solution would merely double the tape consumption without providing significant additional benefits in editing or in post production processing.

TABLE 1 (1976)

	APPROX. SAMPLE FREQUENCY	SAMPLES/ LINE	TOTAL DATA-RATE
YUV	(2+1+1) fsc*	(568, 284, 284)	142 Mbit/s
PAL	2/4 fsc	$567\frac{1}{2}$/1135	71/142 Mbit/s

* PAL sub-carrier frequency

The SMPTE was working towards a proposal based on the digital NTSC signal, sampled at 4 × NTSC sub-carrier frequency. The serial data rate required for this standard is 114 Mbit/s, differing significantly from both of the European proposals (71 and 142 Mbit/s).

These factors led to much reconsideration by users of the PAL system as to their approach to the question of video coding. During 1978, the problems facing PAL users were two-fold:

(i) Was it possible to decode the PAL signal to components without producing unacceptable picture impairments? If it were so, this would allow the adoption of a digital component code on an evolutionary basis.

(ii) If not, was it feasible to consider acceptance of a video standard based on component signals—implying a revolutionary change of studio equipment rather than an evolutionary one?

Some progress has been made on both of these questions, although it must be admitted that significant reservations still remain. A revolutionary change of complete studios to digital working offers certain operational advantages over the 'step-by-step' approach, but would be unlikely to appeal to smaller organisations. The development of very high quality PAL decoders offers a possible solution to this, but these decoders are likely to be costly. The problem of finding an economic and impairment-free path to an optimum component digital code is not yet solved. Nevertheless, users have accepted the need for short-term sacrifices in order to gain the long-term benefits.

By 1979 there was general agreement that a single 625-line digital standard should be developed, and that it be based on separate component signals. The code would be optimised in the Y, R–Y, B–Y domain; compatibility with composite signals being a secondary consideration.

The emergence of a unified approach in Europe was well received by manufacturers of video equipment. The market size would be increased, providing economies of scale, and spreading the costs of equipment development. The question then arose as to whether a 525-line standard built on similar lines would provide further manufacturing advantages.

Just as in PAL countries, users of the NTSC signal

had been working towards a standard based on the digital composite signal. Compatibility of equipment was clearly a major factor in this choice. However, NTSC signals can be decoded with fewer impairments than can PAL or SECAM, and it was possible to consider the co-existence of composite and component codes in 525-line studios. Moreover, it was soon realised that a useful compatibility could be established between 625-line and 525-line component codes, because of the near identity of line frequency. Compatibility of the line parameters would be of significant benefit to manufacturers; certain digital equipments would become dual standard, while many others would contain common sub-units. The SMPTE set up a task force on component coding charged with the responsibility of co-operating with the EBU in the search for a standard component coded television line which would be acceptable world-wide.

Optimisation of the Component Code

The term 'optimisation' can imply a degree of precision which, when making the choice, cannot be achieved in practice. Some of the factors which have been taken into account are:

Picture Quality
Data-rate
Complexity
Signal Processing
Equipment Cost.

It is very difficult to define precisely the relative importance of these factors, although attempts to do so have been made. It was realised at an early stage that picture quality could be related to data-rate, but only through the assumption of a particular sample structure. The bandwidth of analogue television signals is normally defined by using a single parameter corresponding to the horizontal bandwidth. The vertical and temporal bandwidths are implied by the number of lines per field, and the number of fields per second, respectively. For sampled television pictures, the situation is less simple. By adopting sample grids which change in position between adjacent lines, and from frame to frame, it is possible to decrease the sample frequency below the Nyquist minimum without impairing subjective picture quality. Effectively, three-dimensional digital filter options exist which allow an optimum trade-off between horizontal, vertical and temporal resolution. Complicated options of this type were considered and rejected in progressing towards an acceptable digital standard. The problem with proposals of this type is the complexity and cost of using them (not least in the three-dimensional filtering required to recover the baseband signal). However, cost alone was not the deciding factor, because it was clear that the cost of picture manipulation (using field stores) would continue to decrease rapidly within the foreseeable future. The avoidance of complexity as an end in itself was considered as being important.

The sample structures actually adopted are very simple ones (orthogonal in each field), and capable of carrying the full one-dimensional Nyquist bandwidth without recourse to 2-D or 3-D reconstruction filters.

Signal Processing

At the end of 1979 EBU members decided that a proposal based on 12 MHz sampling of the luminance signal, and 4 MHz sampling of each colour difference signal, would be used as a basis for experiments. These frequencies are close to the minimum which can be used (with simple sampling structures) without producing impairments in the Y, R–Y, B–Y domain. Various equipments were constructed, and brought together for test purposes in April 1980.

On tests of picture quality, the (12 : 4 : 4) MHz proposal performed very well.

For a single entry to the digital YUV domain, the picture quality approached that provided by 6 MHz bandwidth RGB signals. There were, however, certain test patterns on which ringing effects could be noticed in the colour-difference channels, but these effects virtually disappeared in subsequent composite encoding operations. These results have since been confirmed in formal subjective tests.

The problems encountered with the (12 : 4 : 4) MHz proposal were not in the area of picture quality, but in signal processing. Two experimental chromakey units had been developed which employed (12 : 4 : 4) MHz signals, and neither of these produced satisfactory results. Despite further development work involving significant improvements in the technique employed, high quality chromakey based on (12 : 4 : 4) MHz processing has yet to be demonstrated.

Data-rate Considerations

Following the tests made in April 1980, a consensus view developed that the standard should be based on a 2 : 1 ratio between the luminance and chrominance sample rates rather than on the 3 : 1 ratio of the (12 : 4 : 4) MHz proposal. Increasing the sample rate in the colour difference channels to create a 2 : 1 ratio would provide significant advantages in the quality of

signal processing (particularly in chromakey), and would offer further advantage of hardware convenience. In view of the inevitable need to cascade digital codecs, and to perform complex digital filtering operations, there was also a commonly held view that a modest increase in the luminance sample-rate would be beneficial. Various 2 : 1 ratio standards, ranging from (12 : 6 : 6) MHz to (14 : 7 : 7) MHz, were considered during 1980/81.

All of these proposals demand data-rates which considerably exceed that of (12 : 4 : 4) MHz signals (see Table 2).

When luminance and colour-difference words are sampled with 8-bit accuracy, (12 : 4 : 4) MHz signals can be carried in a serial bit-stream at 160 Mbit/s. By dropping the line-blanking period, the signal can be made to fit conveniently into the 140 Mbit/s limit available for network links in Europe. Moreover, the 160 Mbit/s figure is close to the maximum which can (at present) be considered for a digital VTR, without employing more than two channels. Before adopting a standard requiring a higher bit-rate, it was necessary to weigh the advantages in signal processing against the possible disadvantages in network distribution and digital recording. New experimental projects were undertaken by EBU and SMPTE members addressing these questions. Joint meetings and demonstrations took place in January and February 1981 in the UK and in the USA.

Development work in support of these meetings largely proved that standards requiring higher data-rates can, without much difficulty, be compressed into the 140 Mbit/s channel. Moreover, the advantages in signal processing afforded by these standards are not necessarily lost in this process[1].

The conclusions in respect of digital recording are somewhat less clear. The recording and replay of perfect pictures at 228 Mbit/s, and with consumption equal to that of existing 1-in. helical machines, was demonstrated. The transfer of tapes between machines, editing in the digital domain and certain 'stunt' modes (including picture-in-shuttle) were also demonstrated. In short, feasibility was proved.

TABLE 2

STANDARD	DATA-RATE
(12 : 4 : 4) MHz	160 Mbit/s
(12 : 6 : 6) MHz	192 Mbit/s
(14 : 7 : 7) MHz	228 Mbit/s

Nevertheless, some doubts persist concerning the cost of producing such machines within the near future.

The Proposed Standard for Digital Studio Interfaces

The intensive work of recent months has resulted in a single proposal which has been submitted to the Technical Committee of the EBU and which has received separate support from members of the SMPTE. This proposal is being considered also by the ABU, OIRT and other interested parties. It is now the only proposal which can lead to a single world-wide standard for the sampling of television lines.

The proposed sampling structures are orthogonal (integer multiples of line frequency); 13.5 MHz for the luminance signal and 6.75 MHz for the colour difference signals. All digital words will be of 8-bits length, resulting in a serial data-rate of 216 Mbit/s.

The proposal is a fairly bold one, placing some trust in the continued advance of digital techniques, particularly in the field of recording. A far-sighted decision of this type minimises the risk that the standard will be overtaken as a result of advancing technology, or by demands for greater quality of flexibility.

It is true that any individual group of users could have developed a standard more suited to its own particular range of activities. The result could then have been a proliferation of incompatible standards, a reduced market size for equipment (with consequent cost increases), and significantly increased problems of programme exchange in Europe. All this has been avoided. Henceforth, studio standards and practices world-wide are on a convergent course.

An Extensible Family of Standards

The code which has been discussed and so far agreed is expected to become the norm in the VTR and post-production areas. It would also be appropriate for use on inter-studio contribution links where further signal processing might be required at the receiving terminal.

What codes should be used in the live studio prior to recording? It is certainly true that signal processing of the highest possible quality demands full bandwidth RGB signals.

What code would be used for distribution links to transmitters? If, on reception, the signal is to be immediately encoded to composite form, is it not wasteful to employ an expensive digital link capable of carrying the studio component code which has requirement in excess of 200 Mbit/s?

Is it possible to foresee digital ENG equipment? Normally, ENG signals would not be subjected to sophisticated signal processing, and could be represented by a code requiring a much lower data-rate.

All of these questions have prompted the SMPTE to introduce the concept of an extensible family of compatible codes, each one designed for a particular application. Care must be taken in the use of this concept. There is a danger that the appearance of too many codes would divide the equipment market into uneconomic sub-units. Nevertheless, alternative codes will be necessary where the studio code cannot satisfy requirements.

A notation has become accepted in which the digital component code is described by a three-number series, these referring to the ratio of the sampling frequencies in the Y, R–Y and B–Y channels respectively. The number representing the sample frequency in the luminance channel of the studio code is arbitrarily chosen as 4. Thus, a code such as (13.5 : 6.75 : 6.75) MHz would be described as a 4 : 2 : 2 code. It is clear that a family of compatible codes may be generated by increasing or decreasing by a factor of 2 the sample frequency in any channel. This may be achieved by a simple interpolation of new samples, or by dropping alternate samples (following a suitable digital filter). Codes such as 8 : 4 : 4, 4 : 4 : 4, 4 : 1 : 1, 2 : 1 : 1 are members of the family. Which members of the family would be useful for tasks which the 4 : 2 : 2 code cannot satisfy?

An RGB Code (4 : 4 : 4)

It has been shown experimentally, and through extensive subjective tests, that chromakey of the highest quality can be achieved only if full bandwidth RGB signals are employed. Today, there is no technique which can give equivalent results by use of bandlimited colour difference channels. The representation of full bandwidth RGB signals would require a code at the 4 : 4 : 4 level.

A digital matrix operation from RGB to YUV is eminently feasible, and has advantages of stability and precision over the established analogue technique. The input RGB signals for such a device demand a 4 : 4 : 4 level code.

Introduction of a (4 : 4 : 4) level code (requiring 300 Mbit/s) upstream of the digital VTR presents no great difficulty. This would allow, in the live studio, very high quality of signal processing. However, in the near future, the recording of such signals on videotape would be very expensive. Nevertheless, certain users are interested in recording these high quality signals, and might be prepared to bear the associated high costs of recording. One possible application would be to take advantage of electronic picture processing in film production. Tape-to-film transfer processes, now available, produce reasonable results when high quality RGB signals are used.

It seems certain that a (4 : 4 : 4) level code would find applications upstream of the VTR, and could be recorded for special applications which justify the costs involved.

An ENG Code (2 : 1 : 1)

It is difficult to imagine digital systems which would challenge the lightweight low-power analogue equipment currently available for ENG. However, there are certain potential problems with analogue equipment of this kind. Lightweight cameras are now approaching the quality of studio cameras, but the same cannot be said of VTRs. The replayed signals are of marginal quality, and will suffer additional impairments if processed in a digital YUV studio centre. The reason is that the signal will suffer two composite decode operations before reaching the viewer. The first decode will occur in the studio centre, and the second in the domestic receiver. (The same will be true of any analogue inputs to a digital YUV studio.) No convincing solution to the 'double decode' problem has yet been found, although one possibility is suggested in the final section of this chapter.

The primary function of ENG is to provide immediacy in news reporting, and the signals are not normally subject to complex processing. Therefore, it would be permissible to consider a digital code providing adequate picture quality, but without the resolution necessary for picture processing (such as chromakey). A 2 : 1 : 1 code might satisfy this requirement.

In constructing a 2 : 1 : 1 code from the 4 : 2 : 2 studio level, it is insufficient to merely halve the sample frequencies without changing the sample structures. Orthogonal structure at the 2 : 1 : 1 level would provide only about 3 MHz of horizontal luminance bandwidth, while leaving the vertical and temporal bandwidths unchanged. A much improved balance can be achieved by suitable filtering (in 2 or 3 dimensions), and of then dropping samples to leave a non-orthogonal structure. Using this technique, recent experiments have shown that the potential quality of digital YUV signals at the 2 : 1 : 1 level (as viewed in the home) is much higher than that of a

typical ENG signal. It remains to be seen whether future lightweight low-power digital recorders can be sufficiently developed as to challenge analogue ENG equipment.

High-resolution Television

This section considers the possible uses of a digital code with quality higher than that of the 4 : 4 : 4 level. There is no demand for high-resolution television as an end in itself. The idea arises from the desire to increase the size of the screen with a consequent increase in viewing angle. It is therefore linked to the development of large flat-screen displays for use in the home, or to film transfer applications.

Existing 525-line or 625-line composite signals provide insufficient quality for viewing distances as low as, say, three times picture height (3H). The defects which can be seen in PAL signals (at 3H) are:

(i) Cross-colour
(ii) Cross-luminance
(iii) Large-area flicker
(iv) Temporal aliasing
(v) Vertical aliasing
(vi) Horizontal resolution
(vii) Vertical resolution.

Cross-colour and cross-luminance are eliminated by use of a separate component transmission system.

Large-area flicker is essentially a function of the display technique. It could be overcome by displaying each field twice within a one-field period, or more efficiently by using a screen which displays information until refreshed.

Temporal aliasing is the effect which causes spoked wheels to appear as if rotating backwards. Increasing the temporal resolution (by increasing field-rate) would not significantly reduce the effect—the right approach is to introduce temporal filtering to remove high temporal frequencies. There is little evidence to suggest that the temporal resolution afforded by the present field rates is insufficient.

Vertical aliasing is likewise the result of an absence of filtering rather than an absence of resolution. This is perhaps the most significant impairment in separate component pictures generated by existing 525/625-line field-interlace systems. Suitable filtering in the vertical direction at the source, and in the display, would remove it. This leaves the question of whether the potential resolution in the horizontal and vertical directions afforded by existing 525/625-line systems is sufficient to permit the use of large screen displays. Available evidence suggests that improvements in picture quality through signal processing will allow a useful increase in the viewing angle without a significant increase in horizontal or vertical bandwidth. However, the signal processing involved will demand a code of higher level than 4 : 4 : 4. The code required is at the 8 : 4 : 4 level, but it would not be formed by a simple doubling of the sample frequencies from the 4 : 2 : 2 code. Rather, the line-frequency itself would be doubled, to provide 4 : 2 : 2 information on all 625 lines in each field (625 lines sequential scan). Spatio-temporal filtering would then be applied, allowing the dropping of alternate lines, thus restoring interlace, and leaving a 4 : 2 : 2 code orthogonal in each field. Such signals could be transmitted (in either analogue or digital form) to a receiver using a 625/525-line separate component channel. The receiver would perform similar spatio-temporal filtering to recover the high quality signals for a 625-line sequential display.

Alternatively, a slightly lower quality would be achieved on the line-interlace display of any conventional receiver if provided with a YUV interface.

Figure 1 shows the range of frequencies for which the conventional line-interlace system produces an alias-free display. Figure 2 shows the improvement which may be gained for the 4 : 2 : 2 line-interlace studio code through the signal processing described[2, 3]. A slight modification of the studio code to a 'field quincunx' arrangement of samples (Fig. 3) allows an alternative baseband alias-free zone as shown in Fig. 4. In this case, the alias-free zone includes horizontal frequencies up to 13.5 MHz, vertical frequencies to 312 cycles per picture height, and temporal frequencies to 25 Hz. However, losses occur in diagonal resolution in comparison with Fig. 2, and it is not yet clear which of these two

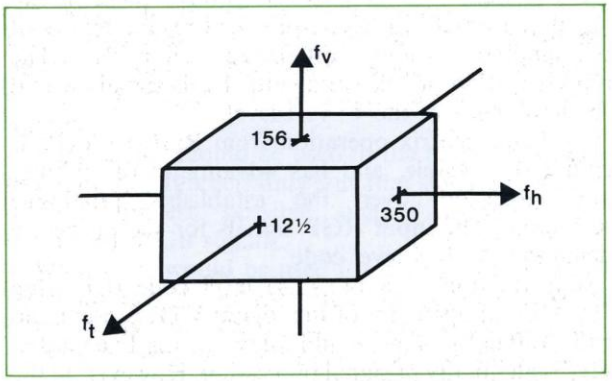


Fig. 1. Frequencies for which the conventional line-interlace produces an alias-free display.

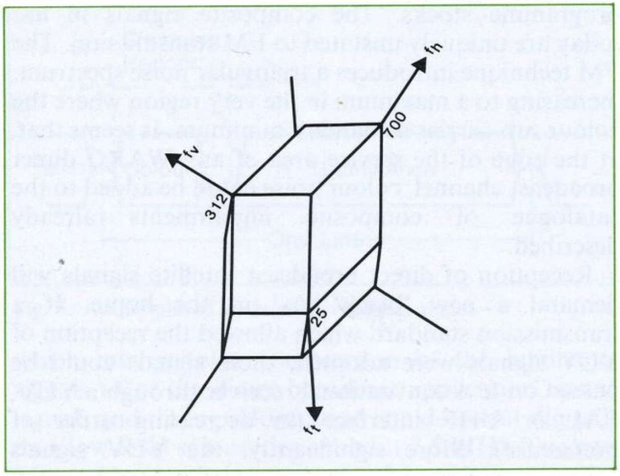


Fig. 2. Improvement gained for 4 : 2 : 2 line interlace studio code.

approaches will give the better subjective results. Both, however, demand the same data-rate as the 4 : 2 : 2 studio code.

A further improvement could be achieved at the same (4 : 2 : 2) data-rate by abandoning line-interlace and employing a 2 : 1 : 1 code and with 625-line sequential-scan. Unfortunately, these signals would be incompatible with the agreed line-interlace 4 : 2 : 2 standard.

Thus, it is evident that future need of a digital studio code at a level higher than 4 : 4 : 4 might arise. That would depend on the development of a large screen domestic display device, and on the appearance of a separate component channel to each receiver. A further application of electronic processing in the film industry seems likely to develop; but the use of studio equipment for this purpose might depend on the aspect ratio required.

New Opportunities

The advent of studio centres operating on a 4 : 2 : 2 digital component code will present new opportunities for improving picture quality by breaking the dominance of composite signals. So long as the composite signal chain remained unbroken, from the source to the transmitter, it was impossible to employ advancing techniques to improve the quality or flexibility of video signals. This would always involve an additional composite decode operation which would cause impairments far outweighing any other advantage which might be gained. The acceptance of a component coding standard in studios will involve temporary difficulties and impairments which will persist until compatibility problems have been solved. These difficulties will themselves accelerate the adoption of separate component signals upstream and downstream of the YUV studio.

Consider, for example, an outside broadcast (OB) operation. Currently, the signals are sent to the studio via analogue composite links. These OB signals (together with any other composite inputs) would suffer a double decode operation before reaching the viewer. All composite signals in use today were invented 20 years ago on the assumption of a single decode in the domestic receiver. They are especially unsuited to applications requiring more than one codec. In fact, for OB links feeding YUV studios in PAL countries, it might actually be better to employ a 625-line NTSC signal to avoid passing twice through PAL. The point is that a new freedom of choice exists for signals upstream of any YUV studio. Of course, the ideal solution would be to use digital YUV links

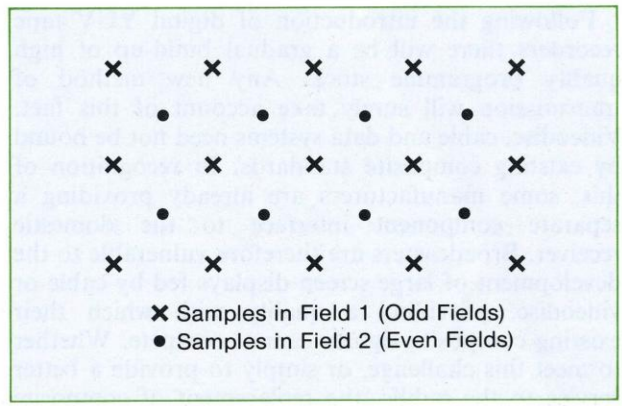


Fig. 3. A field quincunx sample structure.

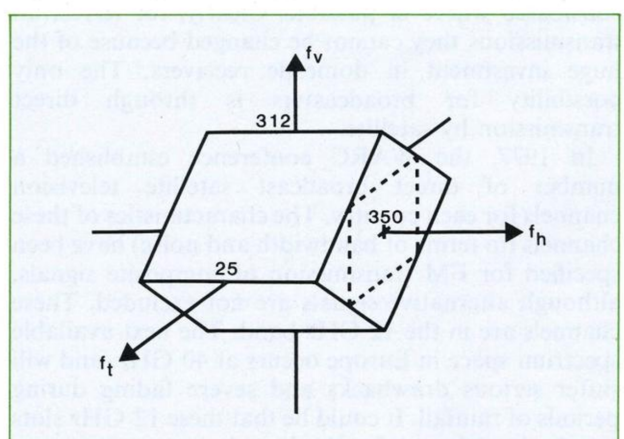


Fig. 4. Baseband alias-free zone.

for OB applications—capable of carrying studio-quality signals. But radio and satellite systems are already in place, and the frequency planning has assumed the use of composite signals. Three choices are therefore available in the immediately foreseeable future.

(i) Continue to use analogue PAL, and suffer the consequences of incompatibility.

(ii) Employ the same radio links to carry digital YUV signals—with heavy bit-rate reduction.

(iii) Employ the same radio links to carry analogue YUV signals.

The second option is a reasonable one in the longer term, but the third option is probably the most practicable in the short term. It requires the development of a new analogue signal based on time multiplexed components.

Also, there are other applications wherein a multiplexed analogue component (MAC) signal might be useful.

Following the introduction of digital YUV tape recorders there will be a gradual build-up of high quality programme stock. Any new method of transmission will surely take account of this fact. Videodisc, cable and data systems need not be bound by existing composite standards. In recognition of this, some manufacturers are already providing a separate component interface to the domestic receiver. Broadcasters are therefore vulnerable to the development of large screen displays fed by cable or videodisc, providing a quality with which their existing composite signals cannot compete. Whether to meet this challenge, or simply to provide a better service to the public, the replacement of composite signals by component coded signals should be considered wherever possible. Clearly, for terrestrial transmissions they cannot be changed because of the huge investment in domestic receivers. The only possibility for broadcasters is through direct transmission by satellite.

In 1977, the WARC conference established a number of direct broadcast satellite television channels for each country. The characteristics of these channels (in terms of bandwidth and noise) have been specified for FM transmission of composite signals, although alternative signals are not excluded. These channels are in the 12 GHz band. The next available spectrum space in Europe occurs at 40 GHz, and will suffer serious drawbacks and severe fading during periods of rainfall. It could be that these 12 GHz slots are the last chance for the broadcasters to improve their signals, and to realise the full potential of future programme stocks. The composite signals in use today are uniquely unsuited to FM transmission. The FM technique introduces a triangular noise spectrum, increasing to a maximum in the very region where the colour sub-carrier demands a minimum. It seems that, at the edge of the service area of any WARC direct broadcast channel, colour noise could be added to the catalogue of composite impairments already described.

Reception of direct broadcast satellite signals will demand a new 'black box' in the home. If a transmission standard which allowed the reception of YUV signals were adopted, these signals could be passed on to a conventional receiver through a YUV, PAL or UHF interface (in decreasing order of preference). More significantly, the YUV signals would become available for a more sophisticated processor/display as described in the section on high-resolution television (separately supplied by videodisc, VCR or fibre optic cable).

What kind of signal would be suitable for high quality direct TV broadcasts? An all digital YUV solution is the most obvious, although the WARC channel characteristics would demand heavy bit-rate reduction, and techniques as yet undemonstrated in a form which would allow for a low-cost receiver option. A new multiplexed analogue component signal is again a more practicable alternative. Basic YUV signals are less sensitive to high-frequency noise than are composite signals. In the region of the colour sub-carrier this difference is as great as 20 dB.

For both the applications which have been discussed (network links and direct broadcast), we need to carry analogue YUV signals within the bandwidth limitations of an FM channel designed for composite PAL. Therefore, the transmission bandwidth available for each component may be lower than that available for the composite signal. Consequently, techniques for maintaining or improving the displayed resolution while decreasing the transmission bandwidth will be needed. These are precisely the characteristics of the three-dimensional signal processing which has been described.

Various possibilities exist for the transmission of separate component signals in an analogue channel. The most promising of these is illustrated in Fig. 5. The two colour difference signals are sent sequentially, line-by-line. They are compressed in time to occupy the line blanking period. The compression process increases the transmission bandwidth of these signals to match that of the luminance channel. (Various non-sequential colour options are possible

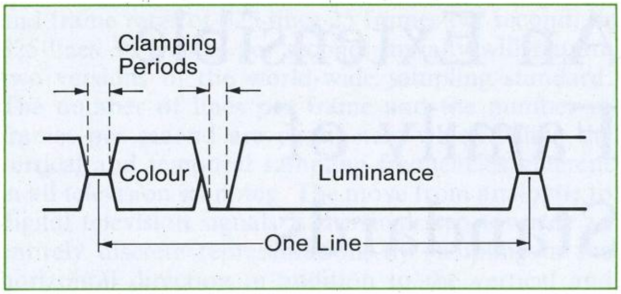


Fig. 5. A multiplexed analogue component signal.

which may involve compression of the luminance signals.) For the direct broadcast application, sound signals and timing information could be sent digitally on a low-level sub-carrier at about 7.5 MHz (this part of the spectrum being too noisy for analogue signals). Decompression in the receiver could be achieved by use of inexpensive CCD technology, the signals being passed on to a conventional receiver. The more expensive signal processing would be optional. MAC signals of the type described have already been considered for experimental high-resolution satellite TV transmission tests in Japan[4].

References

1. 'Bit Rate Reduction to 140 Mbit/s from a Television Component Standard requiring 216–228 Mbit/s', *IBA E & D Interim Report 111/81.*

2. Prof. B. Wendland, University of Dortmund, 'Alternative Lines of Development for Future TV Systems', *1979 Conference of the Fernseh-und Kinotechnischen–Dortmund.*

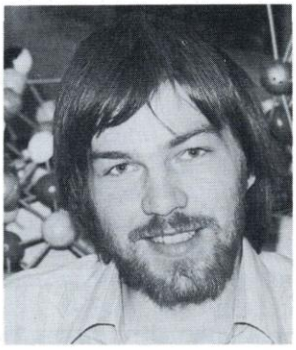
3. Gary Tonge, B.Sc., Ph.D., 'The Sampling of Television Images', *IBA E & D Report No. 112/81.*

4. T. Fujio, T. Komoto & T. Nishizawa, 'High Definition TV System—Signal Standard and Transmission', *NHK Laboratories Note Ser. No. 239* (August 1979).

EDGAR WILSON, B.Sc.(Eng.), ACGI, studied at Imperial College, London, after spending a pre-university year at the BBC Designs Department. After graduation in 1972, he joined the British Post Office to work on systems for digital telecommunications.

From 1974, he has worked in the Independent Broadcasting Authority's Experimental and Development Department, on projects including SABRE and a 60 Mbit/s digital codec for satellite television.

GARY TONGE, B.Sc., Ph.D., graduated from the Electronics Department of Southampton University in 1977. He then moved to the Applied Mathematics Department to research into theoretical limits in energy conversion. He joined the Authority's Video and Colour Section in May 1980 and has since worked on bit-rate reduction and three-dimensional signal processing.



An Extensible Family of Standards

by E. J. Wilson and G. J. Tonge

Synopsis

The preceding chapter in this *IBA Technical Review* deals with factors influencing the choice of a suitable world-wide digital sampling standard for studio quality component television signals. However, one may envisage some important applications for which a studio-quality digital code is not appropriate.

Large-screen displays, for example, may require a code of lower quality. This chapter therefore considers several members of a possible Extensible Family of digital standards encompassing a wide range of requirements. Within this framework the studio standard appears as a 'medium quality' member with other compatible members above and below it.

Introduction

Discussions of digital television coding standards have naturally tended to concentrate on the sampling frequencies and structures necessary to producing pictures to the standards of quality required in the normal studio and post-production environments. This has led to a series of proposals for a studio standard, based on regular orthogonal sampling patterns in each television field having equivalent serial data rates within the range 160 to 228 Mbit/s.

There are certain television applications, such as Electronic News Gathering (ENG), wherein a picture standard of somewhat lower quality would be acceptable. Conversely, we should bear in mind possible future requirements for a television standard superior to the typical studio quality of today.

There is, therefore, the need to consider an extensible family of *compatible* standards including the normal studio one, but with superior and inferior members. Compatibility would assist the designing of equipment which might need to operate at more than one level of quality or to interface between levels.

It is theoretically possible to interface between any arbitrary pair of coding standards, but the design problems are eased if the relationship between the sampling rates is simple. The second part of this chapter considers some possible members of an extensible family of compatible standards and describes results obtained from practical demonstrations of some such members.

Basic Studio Standard

At the time of writing, there appears a likelihood that world-wide agreement on the digital sampling frequency for studio television use will be achieved. Of course, the sampling frequency is but one of many parameters which define a television system, and the major parameters are those inherited from the present analogue systems. In particular, the line structures

and frame rates of 625-lines 25 frames per second, or 525-lines 30 frames per second, initially will require two versions of the world-wide sampling standard. The number of lines per frame and the number of frames per second are parameters which define the vertical and temporal sampling frequencies inherent in all television scanning. The move from analogue to digital television signals is the final step towards an entirely discrete representation, by sampling in the horizontal direction in addition to the vertical and temporal sampling already applied.

If horizontal sampling operates with an orthogonal horizontal/vertical sample structure, then each frame consists of an identical pattern of samples. This greatly simplifies some signal processing techniques such as picture-store design. The discussions about sampling standards for studio applications have taken account of this and have led to proposals based upon orthogonal sampling.

The likely parameters of a future standard for digital studio equipment are shown in Table 1.

For ease of comparison with other members of a standards hierarchy, a short form of nomenclature has been adopted which labels the studio standard as 4 : 2 : 2 (Y : U : V) code, these figures referring to the ratio of sample frequencies in the three channels. The sampling structures for luminance and chrominance are illustrated in Figs. 1 and 2. Possible members of a compatible family, shown in similar notation, are listed in Table 2.

A 625-line per Field Code (8 : 4 : 4)

The European analogue PAL and SECAM television systems are already sampled two dimensionally, i.e. vertically and temporally. The line-interlace scan broadcast system may be seen as one method of downsampling from a theoretical source standard

TABLE 1

Luminance Sampling Frequency	13.5 MHz Line Locked (864 Samples Per Line)
Luminance-to-Chrominance sampling ratio	4 : 2 : 2 (Y : U : V)
Sampling structure	Horizontal/vertical orthogonal line interlaced between alternate fields

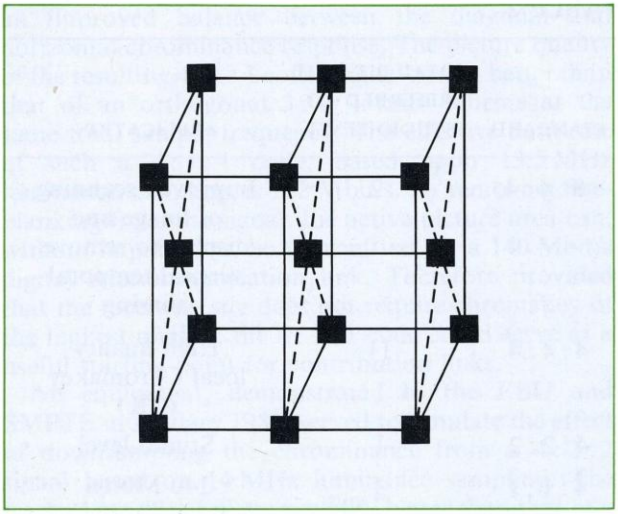


Fig. 1. 4 : 2 : 2 luminance sampling structure with interlace.

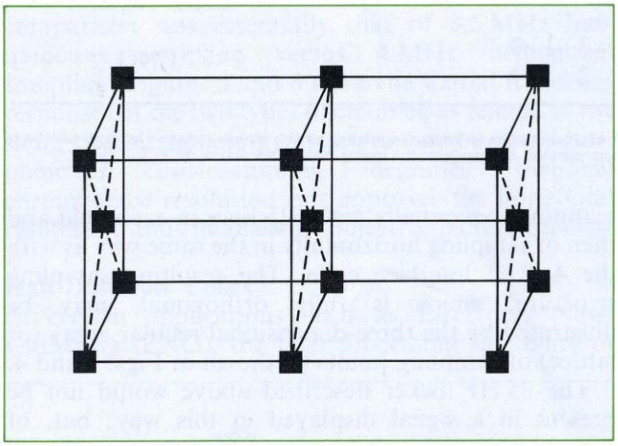


Fig. 2. 4 : 2 : 2 chrominance sampling structure with interlace.

with sequential scanning at 625-lines per field, 50 fields per second. The interlace technique retains most of the resolution available from the 625-line sequential standard, but at a saving of half the transmission bandwidth.

Interlaced scanning, as now used universally for broadcast television, suffers certain limitations which are becoming more noticeable with high-resolution component television signal origination. Aliasing, which is visible as a 25 Hz flicker on high vertical spatial frequencies, arises from the scanning procedure. For applications involving large-screen displays (on which this effect becomes more objectionable) a non-interlace option in the digital hierarchy is desirable. One possibility is that of

TABLE 2

STANDARD	TOTAL BIT-RATE REFERRED TO STUDIO LEVEL	APPLICATION
8 : 4 : 4	2	Improved scanning of image and display to remove vertical/temporal aliasing
4 : 4 : 4	$1\frac{1}{2}$	High quality local Chromakey (CK)
4 : 2 : 2	1	Studio level
4 : 1 : 1	$\frac{3}{4}$	140 Mbit/s transmission* where no downstream CK is required
3 : 1 : 0	$\frac{1}{2}$	ENG
2 : 1 : 1	$\frac{1}{2}$	

* Alternative codes for transmission which employ bit-rate reduction methods have been demonstrated, and can provide a quality higher than that of the 4 : 1 : 1 code[1].

scanning sequentially with 625-lines in each field and then of sampling horizontally in the same way as with the 4 : 2 : 2 interlace code. The resulting sampling structure, which is fully orthogonal, may be illustrated by the three-dimensional regular array (or lattice) of sampling points as shown in Figs. 3 and 4.

The 25 Hz flicker described above would not be present in a signal displayed in this way; but, of course, the 50 Hz field sampling rate would still produce the much less objectionable 50 Hz flicker.

It is possible to achieve many of the benefits of the 8 : 4 : 4 code without need of either transmitting or storing the signal at the associated double bit-rate. By pre-filtering the signal in the vertical/temporal frequency domain, and subsequently omitting alternate television lines, a new 4 : 2 : 2 interlaced signal is produced. This signal could be processed or stored at the studio bit-rate, and finally resampled with post-filtering, for display on a 625-line per field 8 : 4 : 4 level monitor. The 25 Hz alias associated with conventional interlace would be eliminated by the downsampling pre-filter and post-filter. The pre-filtered 4 : 2 : 2 signal would be directly compatible for viewing on displays at the interlace 4 : 2 : 2 level or at lower levels. Should any particular application require the full unfiltered resolution of the 8 : 4 : 4 code, intermediate downsampling would be avoided. A further discussion of options for the 8 : 4 : 4 code, and a description of some possible practical filters for 8 : 4 : 4 – 4 : 2 : 2 – 8 : 4 : 4 conversion processing are given in Ref. 2.

A High Quality Chromakey Code 4 : 4 : 4

One of the major requirements for the studio 4 : 2 : 2 code is that the chrominance bandwidths afforded by the 4 : 2 luminance-to-chrominance ratio should provide excellent chromakeying quality[3]. Subjective tests of the chrominance bandlimiting filters used with 6–7 MHz sampling frequencies have shown 'just-perceptible' impairment of chromakeyed pictures.

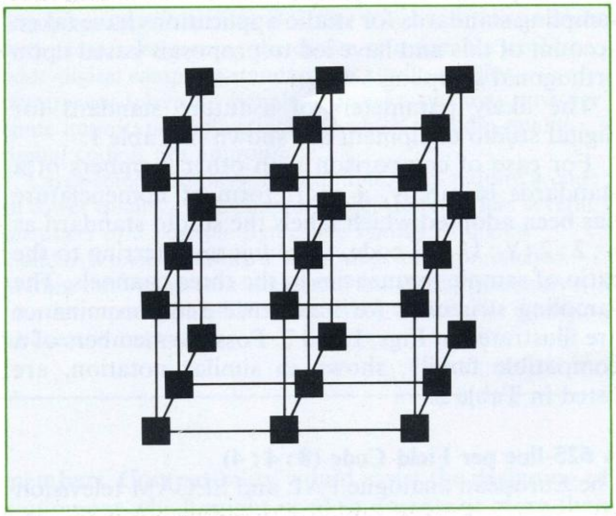


Fig. 3. 8 : 4 : 4 luminance sampling structure.

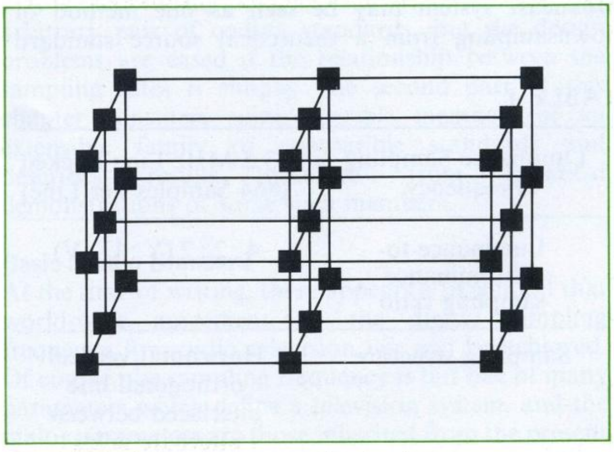


Fig. 4. 8 : 4 : 4 chrominance sampling structure.

Increase of the chrominance sampling frequency to that of luminance (i.e. 4 : 4 : 4 ratio) produced slightly better results; and these were exactly similar to those obtained by use of direct RGB analogue camera feeds. There are certain circumstances in which very high quality chromakey performance is required; and, for this purpose, a 4 : 4 : 4 code would be necessary.

After use for chromakey or other signal processing, the 4 : 4 : 4 code can be converted to a signal of studio level. This requires horizontal filtering of the chrominance signals and downsampling by alternate sample omission.

A Reduced Chrominance Bandwidth Code 4 : 1 : 1

If a studio code were chosen for picture quality requirements alone, the chrominance bandwidth afforded by a 4 : 2 : 2 ratio code might be considered excessive. In fact, early work[4] showed that 3 : 1 : 1 ratio standards were generally acceptable in this regard. However, a 4 : 1 : 1 orthogonally sampled signal at the same total data-rate would probably be considered inadequate; but, when a 4 : 1 : 1 signal is generated from the 4 : 2 : 2 studio code, chrominance downsampling need not use simple one-dimensional filtering and sample omission. Use of two-dimensional filtering followed by line quincunx downsampling produces a line-interleaved sample pattern; and thereby, diagonal chrominance spatial bandwidth may be traded for pure horizontal and/or pure vertical bandwidth. The technique is known as sub-Nyquist filtering (see Appendix); and it leads to an improved balance between the diagonal and horizontal chrominance response. The picture quality of the resulting 4 : 1 : 1 coded signal can be better than that of an orthogonal 3 : 1 : 1 ratio scheme at the same total sample frequency. The effective data-rate of such a 4 : 1 : 1 code, based upon 13.5 MHz luminance sampling, is 162 Mbit/s. By removing line-blanking from the signal, the active picture area can, without impairment, be transmitted via a 140 Mbit/s digital telecommunication link. Therefore provided that the receiving site does not require chromakey of the highest quality, the 4 : 1 : 1 code could serve as a useful starting point for contribution links.

An equipment, demonstrated to the EBU and SMPTE in January 1981, served to simulate the effect of downsampling the chrominance from a 4 : 2 : 2 signal based on 14 MHz luminance sampling. The method produced picture quality better than that of a 3 : 1 : 1 signal based on 12 MHz luminance sampling for a similar overall data rate. As to chrominance, the comparison was essentially that of 3.5 MHz line-quincunx sampling versus 4 MHz orthogonal sampling. Figures 5 and 6 show the spatial frequency responses of the two types of channel, as related to the same overall sampling rates. As can be seen, line-quincunx downsampling degrades diagonal chrominance resolution but improves the horizontal resolution, thus increasing subjective picture quality.

Half Data-rate Codes

In certain applications such as Electronic News Gathering (ENG), the high processing capability of

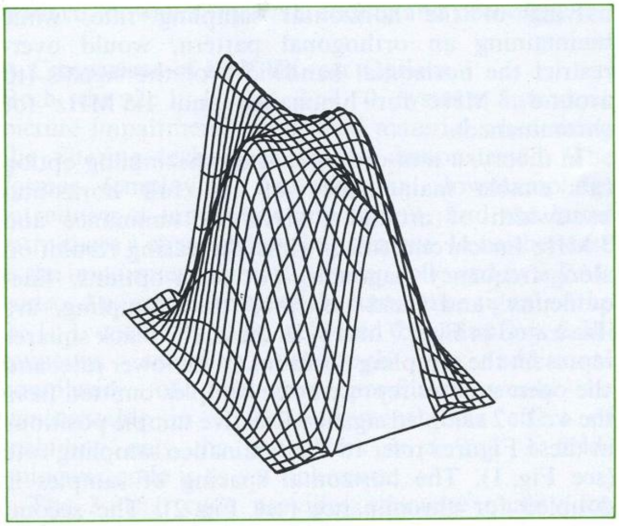


Fig. 5. Line-quincunx downsampling pre-filter.

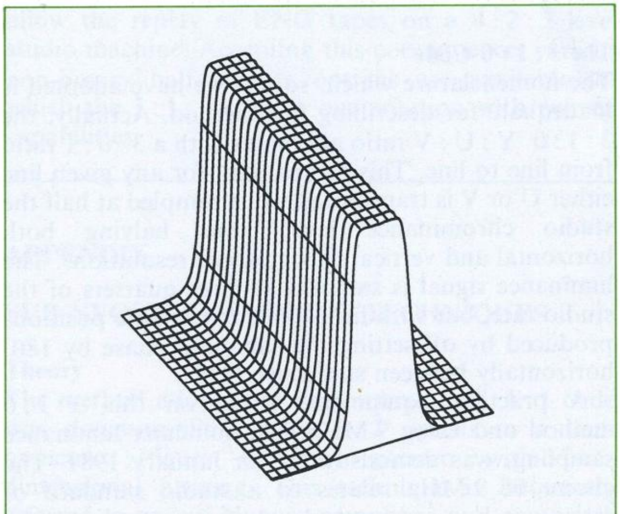


Fig. 6. Orthogonal downsampling pre-filter.

the 4 : 2 : 2 studio code is unnecessary; and it may be sacrificed in order to simplify the signal originating and recording equipment. In such cases, a half bit-rate option would be attractive, provided that it can produce pictures of broadcast quality. A quality comparable with (or possibly better than) that of composite television pictures would be demanded. Of the many possibilities for halving the overall data-rate, two methods were demonstrated during EBU and SMPTE technical viewing tests, and will here be described.

The two methods adopt quite different design philosophies. The first (here termed as a 3 : 1 : 0 code) samples directly the luminance scene at three-quarters of the studio luminance rate. The major data-rate saving is effected in the chrominance signals where alternate lines of U or V are dropped after sampling at half the (4 : 2 : 2) studio rate. Very little processing at the camera is required to produce the 3 : 1 : 0 signal; and it can be recorded directly in that form. However, considerable processing is required to reconstruct the 4 : 2 : 2 level code in the studio. An optimum reconstruction requires three-dimensional processing; but, reasonably good results can be achieved by alternative means. The second, a 2 : 1 : 1 method, uses two-dimensional sub-Nyquist filtering of all signals at the camera source and for reconstruction of the 4 : 2 : 2 code. Recording or transmission can take place at the half bit-rate (2 : 1 : 1) level. The physical requirements for two-dimensional filtering are not severe, and, with use of future VLSI techniques, might represent little burden on the camera-recorder combination.

The 3 : 1 : 0 Code

The nomenclature which, so far, we have adopted is inadequate for describing this method. Actually, the 3 : 1 : 0 Y : U : V ratio alternates with a 3 : 0 : 1 ratio from line to line. This means that, for any given line either U or V is transmitted while sampled at half the studio chrominance level, thus halving both horizontal and vertical chrominance resolutions. The luminance signal is sampled at three-quarters of the studio rate, but with field-quincunx sample positions produced by off-setting the sampling phase by 180° horizontally between successive fields.

A practical equipment, based on this 3 : 1 : 0 method and using 9 MHz field-quincunx luminance sampling, was demonstrated* in January 1981. The choice of 9 MHz relates to a studio standard of 12 MHz luminance sampling, but, for the demonstration, no attempt was made to digitally resample the 3 : 1 : 0 signal to the equivalent 4 : 2 : 2 level. Instead, the interface was effectively an analogue one; and it plainly showed the ultimate resolution derivable from such conversion. Optimum luminance reconstruction (to the 4 level) requires fairly complex three-dimensional filtering; but the need of only one studio resampling equipment handling several ENG sources might, financially, justify the method. Alternatively, the field-quincunx luminance signal can be presented without post-filtering, leaving the temporal response of the eye to partially cancel the alias products in the signal. The main picture impairment occurs in the chrominance channels due to the loss of alternate lines in each of the U and V signals. If lacking pre-filtering and post-filtering, this process gives rise to aliasing on chrominance horizontals. Conversely, inclusion of such filters serves to reduce to a marginal level the vertical chrominance resolution. Probably, the best balance is to be achieved by excluding the pre-filter (thereby maintaining simplicity in the ENG camera), and including a post-filter/interpolator. The 3 : 1 : 0 method requires only a minimum of signal processing at the camera/recorder. It might, therefore, satisfy ENG needs.

A 2 : 1 : 1 Code

This code is generated by downsampling, by a factor of two, each of the 4 : 2 : 2 studio signals, and with appropriate filtering. The many possible options for effecting this are fully explored in Ref. 2. Simple halving of the horizontal sampling rate, while maintaining an orthogonal pattern, would over-restrict the horizontal bandwidth of the signals (to around 3 MHz for luminance and 1.5 MHz for chrominance).

In theory, a non-orthogonal downsampling option can enable maintenance of the full horizontal bandwidth of around 6 MHz for luminance and 3 MHz for chrominance, while decreasing resolution along frequency diagonals. Two such options, 'line-quincunx' and 'field-quincunx' downsampling, are illustrated in Figs. 7 and 8. In these, the black squares represent the sampling positions at the lower rate, and the open squares represent the samples omitted from the 4 : 2 : 2 sampled signals. Relative sample positions in these Figures refer to the luminance sampling rate (see Fig. 1). The horizontal spacing of samples is doubled for chrominance (see Fig. 2). The second option (field-quincunx) seems optimal for luminance

* Equipment designed by IRT, Germany.

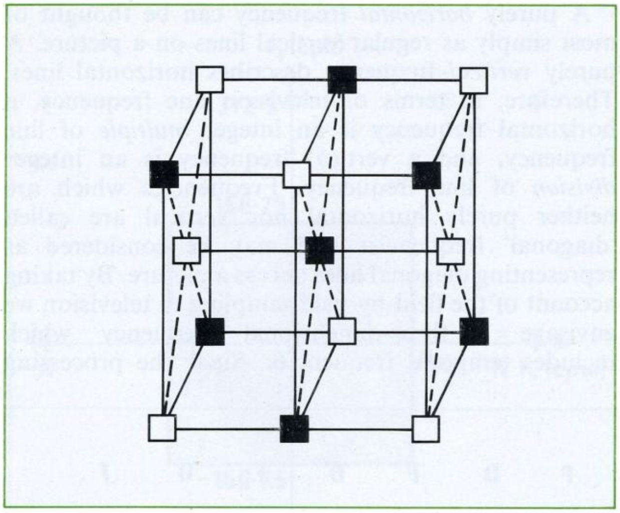


Fig. 7. Line-quincunx downsampling.

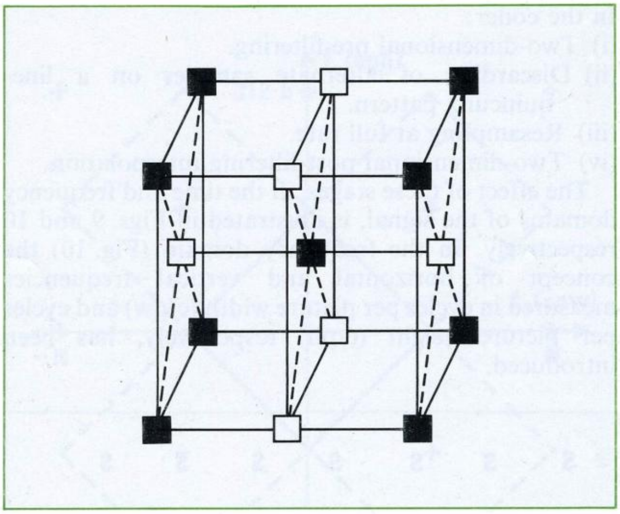


Fig. 8. Field-quincunx downsampling.

because of its symmetry in the horizontal, vertical and temporal directions. However, the filtering required to implement this case involves shift register storage of several television fields; which, on consideration of current size and cost, might be deemed uneconomical. Therefore, the practical method which was demonstrated was restricted to simulating two-dimensional filtering for line-quincunx downsampling of both luminance and chrominance using a maximum of three television lines (see Appendix).

Using this method at the (14 : 7 : 7) MHz sampling rate a perceptible loss of luminance diagonal resolution is introduced. This impairment is largely masked, however, by PAL coding and decoding.

A Comparison of Half Bit-rate Options

Both the 2 : 1 : 1 and 3 : 1 : 0 systems introduce picture impairments on critical material when using the filtering techniques so far demonstrated. The former (employing line-quincunx downsampling) introduces a luminance impairment, and the latter introduces a chrominance impairment. In neither case is the impairment significant when followed by a PAL codec. An improvement to the luminance signal of the 2 : 1 : 1 code might be achieved by adopting field-quincunx sampling, although the additional complexity of field-store pre-filters would be unacceptable in ENG equipment. The effect of including only the post-filters in a 2 : 1 : 1 field-quincunx code is as yet unknown.

The 3 : 1 : 0 code provides adequate quality for ENG equipment, and without need of pre-filtering. This would lead to very simple processing in ENG cameras. The 2 : 1 : 1 code requires line-store pre-filtering, but offers the possibility of a modular approach in a family of digital recorders which includes the 4 : 2 : 2 level studio machine.

Should LSI technology advance to the point where a line-store pre-filter can be included in ENG equipment, or should achievement of a 2 : 1 : 1 code of adequate quality and with simplified pre-filtering prove possible, then 2 : 1 : 1 would be a favourable choice. If not, then the concept of modular VTR design must be abandoned; and it might prove impossible or uneconomical to consider designs which allow the replay of ENG tapes on a 4 : 2 : 2 level studio machine. Accepting this consequence, various non-binary half bit-rate options are available, of which the 3 : 1 : 0 code is one solution with proven capabilities.

APPENDIX

SUB-NYQUIST SAMPLING TECHNIQUES

Theory

The method adopted to demonstrate a 2 : 1 : 1 code was downsampling the 4 : 2 : 2 signals in a line-quincunx (figure of five) pattern with two-dimensional filtering. This technique is frequently referred to as sub-Nyquist sampling[5] and comprises the following stages:

In the coder:

(i) Two-dimensional pre-filtering.

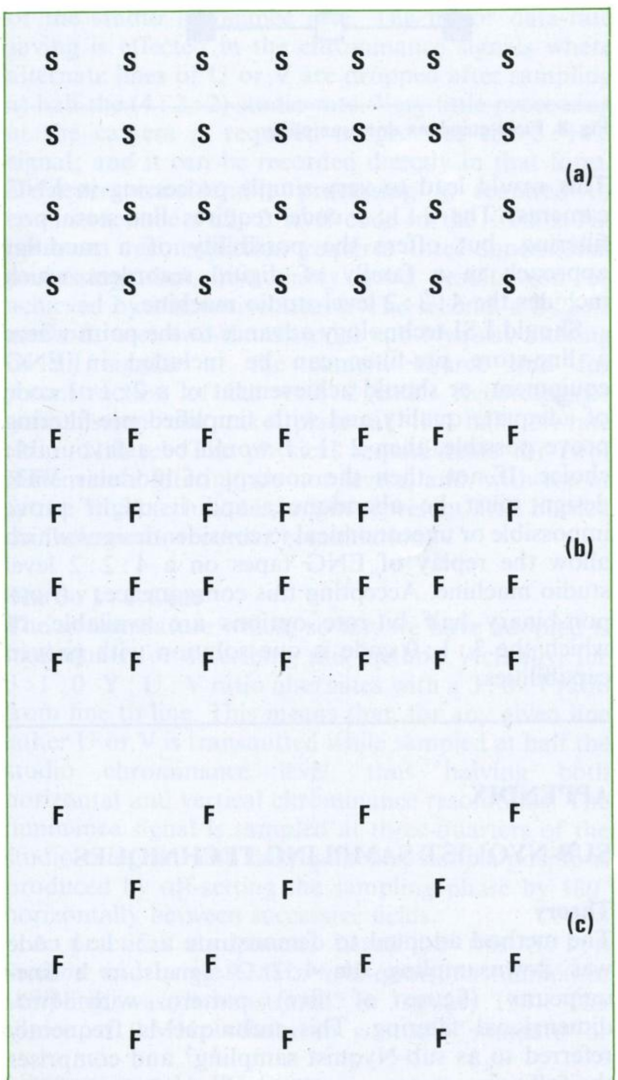
(ii) Discarding of alternate samples on a line-quincunx pattern.

(iii) Resampling at full rate.

(iv) Two-dimensional post-filtering/interpolation.

The effect of these stages, in the time and frequency domains of the signal, is illustrated in Figs. 9 and 10 respectively. In the frequency domain (Fig. 10) the concept of horizontal and vertical frequencies measured in cycles per picture width (c/pw) and cycles per picture height (c/ph) respectively, has been introduced.

A purely *horizontal* frequency can be thought of most simply as regular vertical lines on a picture. A purely *vertical* frequency describes horizontal lines. Therefore, in terms of television line frequency, a horizontal frequency is an integer *multiple* of line frequency, and a vertical frequency is an integer *division* of line frequency. Frequencies which are neither purely horizontal nor vertical are called 'diagonal' frequencies and may be considered as representing diagonal lines across a picture. By taking account of the field-by-field sampling in television we envisage a three-dimensional frequency which includes temporal frequencies. Since the processing



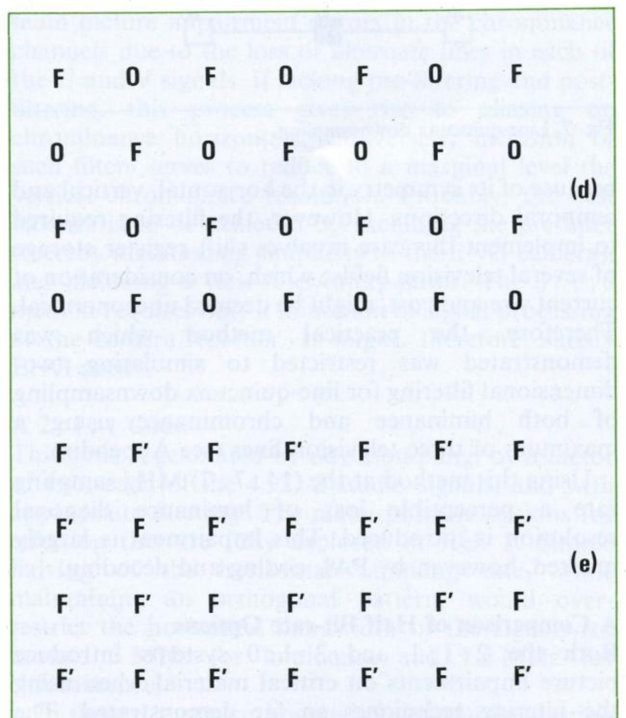


Fig. 9. Sub-Nyquist sampling in the time domain.
(a) Relative sample positions for orthogonal sampling on a single field. There are N samples on each line.
(b) Stage (i) of the sub-Nyquist process—filtering F.
(c) Stage (ii)—discarding of alternate samples in a line-quincunx pattern.
(d) Stage (iii)—resampling at full rate inserts zeros in missing sample positions.
(e) Stage (iv)—post-filtering replaces zero samples by interpolated values F′.

Horizontal frequencies are often quoted in units of MHz. Frequencies in this form may be expressed as cycles per picture width by dividing by the appropriate line frequency.

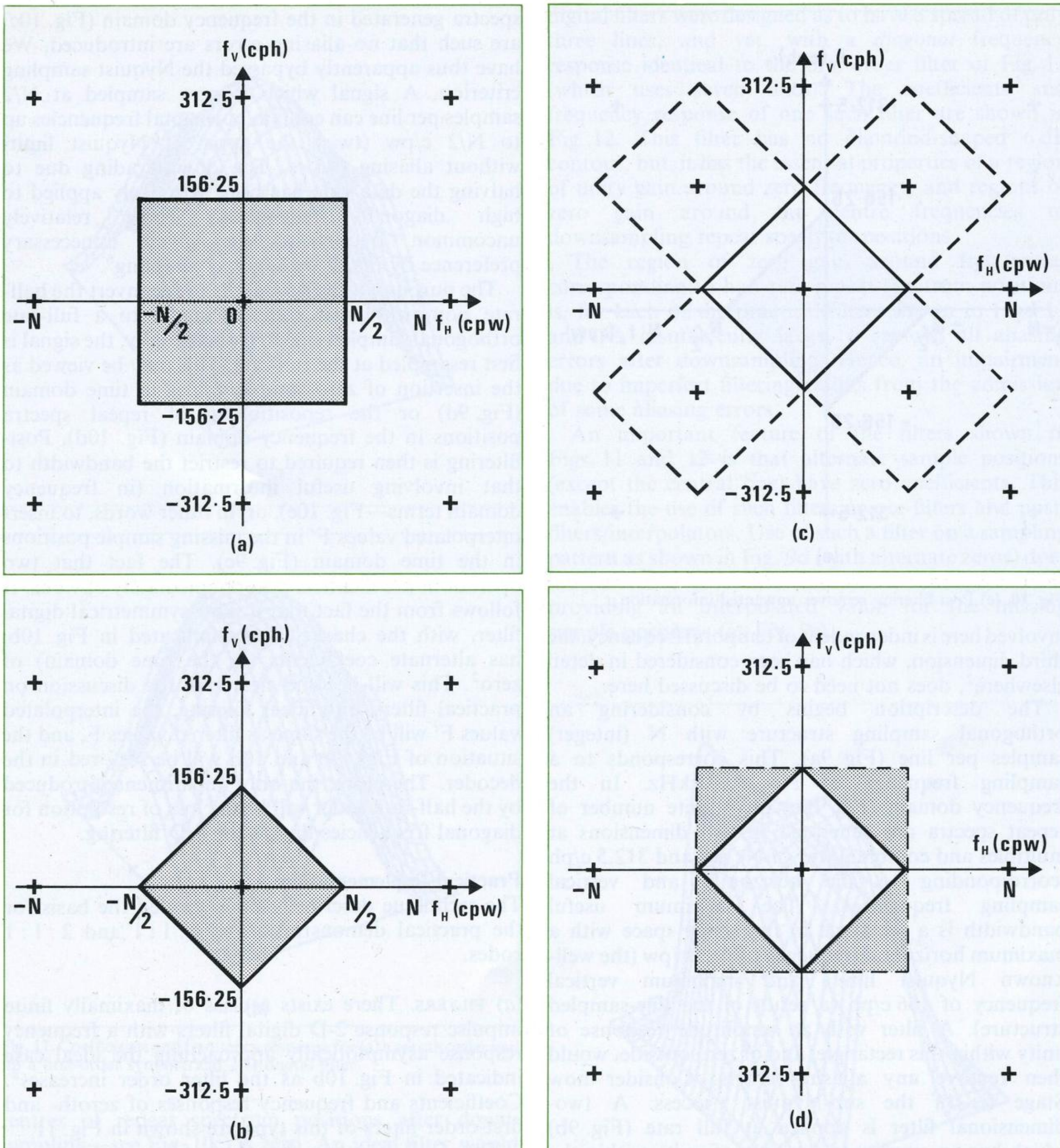


Fig. 10. Sub-Nyquist sampling in the frequency domain. Figures show repeat spectrum positions (black crosses) on a grid of horizontal frequencies (FH) and vertical frequencies (FV).
(a) Shaded region denotes useful bandwidth for an orthogonally sampled signal with N samples per line.
(b) A 2-D filter is applied which limits the bandwidth further reducing diagonal resolution.
(c) Omission of alternate samples in a line-quincunx pattern gives rise to further repeat spectra. These are such that no aliasing errors are introduced (diamond filter shapes do not overlap).
(d) Resampling at full rate removes some repeat spectra positions. Information within diamond shape is useful; that within dotted region is not.

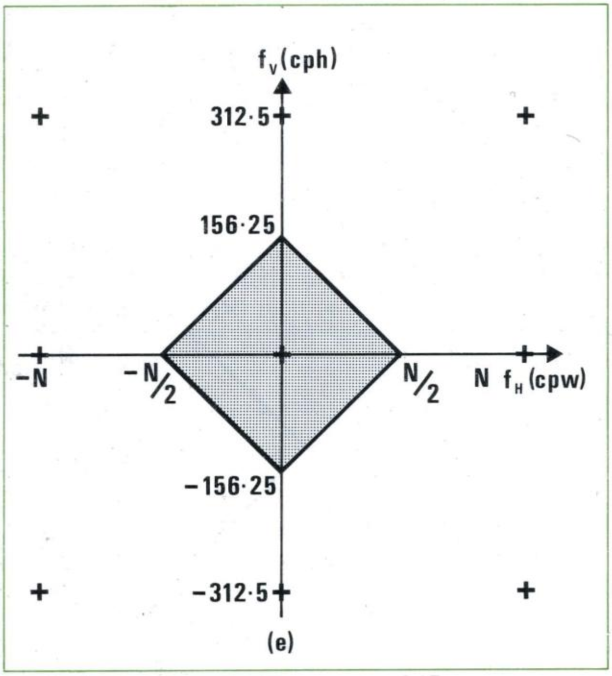


Fig. 10. (e) Post-filtering removes unwanted information.

involved here is independent of temporal frequency, the third dimension, which has been considered in detail elsewhere[2], does not need to be discussed here.

The description begins by considering an orthogonal sampling structure with N (integer) samples per line (Fig. 9a). This corresponds to a sampling frequency of $N \times 15.625$ kHz. In the frequency domain (Fig. 10a) an infinite number of repeat spectra are generated in two dimensions at multiples and combinations of Nc/pw and 312.5 c/ph (corresponding to the horizontal and vertical sampling frequencies). The maximum useful bandwidth is a rectangle in frequency space with a maximum horizontal frequency of N/2 c/pw (the well-known Nyquist limit) and maximum vertical frequency of 156 c/ph (a result of the line-sampled structure). A filter with an amplitude response of unity within this rectangle, and of zero outside, would then remove any aliasing errors. Consider now Stage (i) of the sub-Nyquist process. A two-dimensional filter is applied at full rate (Fig. 9b) which has an amplitude response of unity within the shaded area in Fig. 10b, and of zero outside. The effect of this filter is to limit *diagonal resolution* while maintaining horizontal and vertical resolution. If alternate samples are then omitted (Stage (ii)) in a line-quincunx pattern (Fig. 9c) the additional repeat spectra generated in the frequency domain (Fig. 10c) are such that no aliasing errors are introduced. We have thus apparently bypassed the Nyquist sampling criterion. A signal which is now sampled at N/2 samples per line can contain horizontal frequencies up to N/2 c/pw (twice the apparent Nyquist limit) without aliasing errors. The down-grading due to halving the data rate has been selectively applied to high diagonal frequencies. These relatively uncommon frequencies are given unnecessary preference by initial orthogonal sampling[6].

The purpose of the decoder is to reconvert the half-rate non-orthogonal sampled signal to a full-rate orthogonal sampled signal. Conceptually, the signal is first resampled at the full rate. This may be viewed as the insertion of zero sample values in time domain (Fig. 9d) or the repositioning of repeat spectra positions in the frequency domain (Fig. 10d). Post-filtering is then required to restrict the bandwidth to that involving useful information (in frequency domain terms—Fig. 10e), or, in other words, to insert interpolated values F′ in the missing sample positions in the time domain (Fig. 9c). The fact that two operations are identical is not immediately obvious. It follows from the fact that a skew symmetrical digital filter, with the characteristics indicated in Fig. 10b, has alternate coefficients (in the time domain) of zero[2]. This will become clearer in the discussion on practical filters. For ideal filtering, the interpolated values F′ will be the same as filtered values F, and the situation of Figs. 9b and 10b will be restored in the decoder. Therefore, the only impairment introduced by the half-rate coder will be the loss of resolution for diagonal frequencies due to the 2-D filtering.

Practical Implementation

The technique described above formed the basis for the practical demonstration of 4 : 1 : 1 and 2 : 1 : 1 codes.

(*a*) FILTERS. There exists a class of maximally finite impulse response 2-D digital filters with a frequency response asymptotically approaching the ideal case indicated in Fig. 10b as the filter order increases[2]. Coefficients and frequency responses of zeroth- and first-order filters of this type are shown in Fig. 11.

The 6 dB contour for these filters corresponds exactly to the ideal filter contour shown in Figs. 10b-e. The amplitude response in both cases for zero horizontal and vertical frequencies (the centre of the diamond-shaped region) is unity. The response at the four corners of the plot (which correspond to the

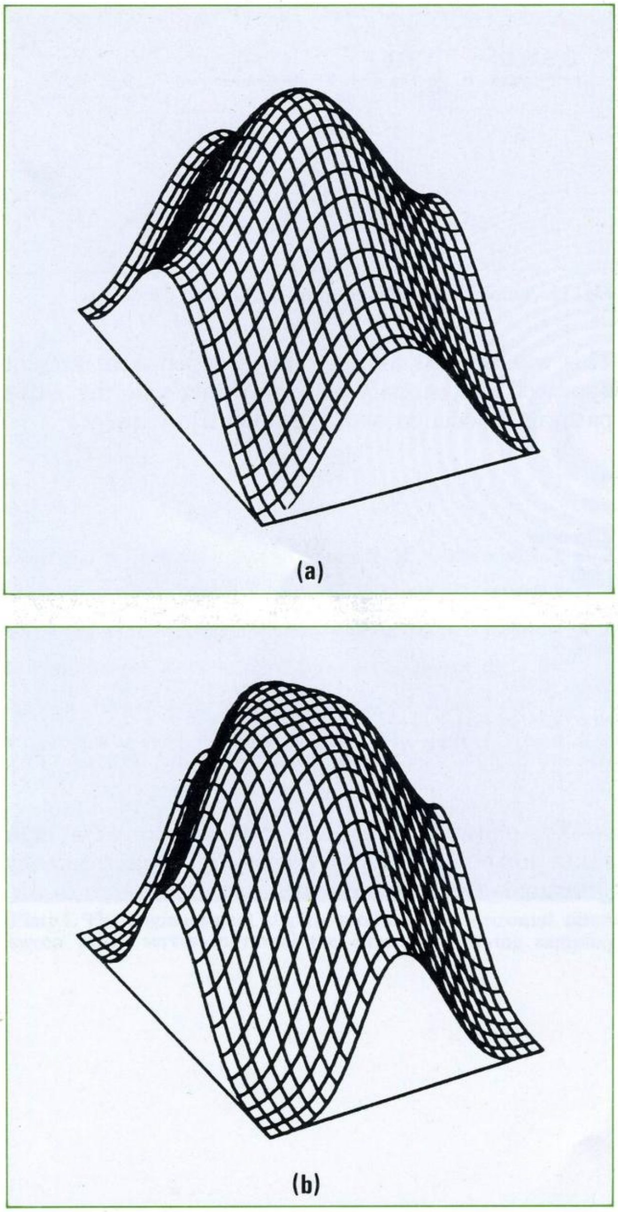


Fig. 11. Coefficients and frequency response for (a) a zeroth order and (b) a first-order symmetrical 2-D digital filter.

centres of repeat spectrum positions after downsampling—see Fig. 10c) is zero. An ideal filter would have a response of unity within the whole diamond-shaped region and of zero outside that region. Economy of hardware prompted the decision to limit to three the number of lines required for filter-processing. Because of this restriction, several 2-D digital filters were designed as to have a spread of only three lines, and yet, with a *diagonal* frequency response identical to the first-order filter of Fig. 11 (which uses seven lines). The coefficients and frequency response of one such filter are shown in Fig. 12. This filter has no diamond-shaped 6 dB contour, but, it has the essential properties of a region of unity gain around zero frequency, and regions of zero gain around the centre frequencies of downsampling repeat spectrum positions.

The region of zero gain around frequencies corresponding to half-rate repeat-spectrum positions is, for each of the practical filters shown in Figs. 11 and 12, insufficiently large to remove all aliasing errors after downsampling. Hence, an impairment due to imperfect filtering results from the admission of some aliasing errors.

An important feature of the filters shown in Figs. 11 and 12 is that alternate sample positions (except the central one) have zero coefficients. This enables the use of such filters as pre-filters and post-filters/interpolators. Use of such a filter on a sampling pattern as shown in Fig. 9d (with alternate zeros) does indeed leave alternate samples unchanged while providing an interpolated value for the missing sample positions (as Fig. 9e).

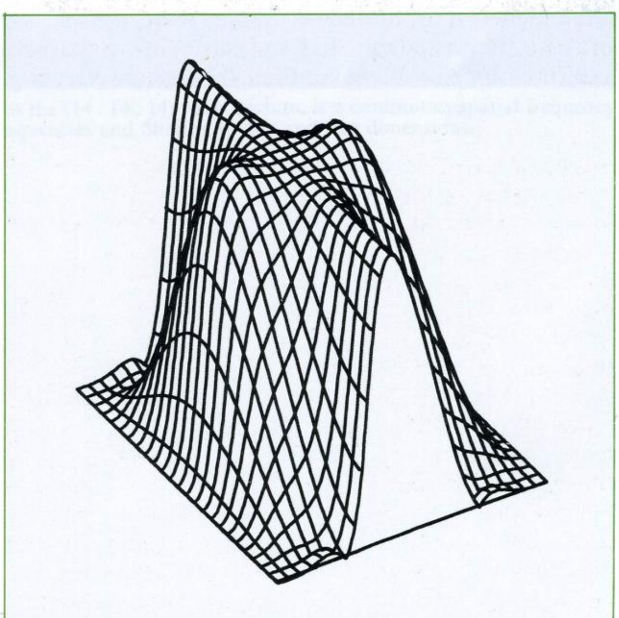


Fig. 12. Coefficients and frequency response of a three-line 2-D digital filter with first-order diagonal characteristics.

(*b*) SUB-NYQUIST SYSTEM. The above feature of the filters useful in a sub-Nyquist system makes possible a precise simulation of sub-Nyquist coding and decoding without need of actually downsampling to half-rate. This is illustrated in the block diagram of Fig. 13, wherein a video signal (Y, U or V) is sampled, thereby yielding values S, S, S, S in a continuous stream. The signal is then passed through a 2-D digital filter yielding a stream of F, F, F, F. A second filter has the same coefficients as the first, except that it has a central coefficient of zero. Passing of the signal through this second filter results in a stream of values F′, F′, F′, F′, where the F′ are the same as the alternate interpolated values in a true sub-Nyquist decoder (see discussion above and Fig. 9e). Switching between the stream F′, F′, F′, F′ and the stream F, F, F, F (suitably delayed by the equivalent filter delay τ) for alternate samples then yields a stream F, F′, F, F′ precisely as required to simulate sub-Nyquist coding and decoding. Any signal sampled with an even number of samples per line requires that the sample selection switch be reset for each line. Thereby, the F and F′ samples in two-dimensions form a line-quincunx pattern as shown in Fig. 9e.

The line-quincunx sampling structure associated with sub-Nyquist sampling repeats every four fields. For some applications a two-field repeating sequence is more useful; and so, the sub-Nyquist simulation equipment incorporated a frame reset option. In practice, it was found that for stationary pictures a subjective *improvement* resulted from such resetting.

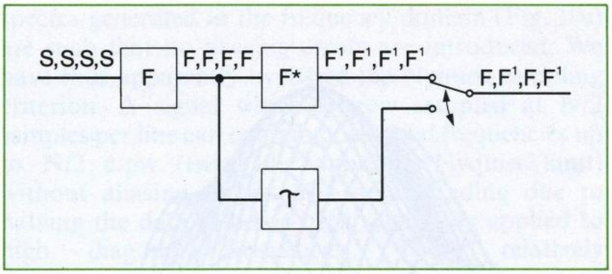


Fig. 13. Simulation of sub-Nyquist coding and decoding.

This was because aliasing errors which were present appeared as stationary patterns rather than the active patterns associated with the four-field sequence.

References

1. E. J. Wilson and P. R. Carmen, 'Bit-rate Reduction for 140 Mbit/s Links'.
2. G. J. Tonge, 'The Sampling of Television Images', *IBA E & D Report 112/81.*
3. R. Rawlings and N. J. Seth-Smith, 'Chromakey in Future Studio Systems'.
4. 'CCIR Draft Report' 629–1 (Mod I) 'CCIR Draft Report' (1978–82 11/14) (EBU).
5. K. H. Barratt and K. Lucas, 'An Introduction to sub-Nyquist Sampling', *IBA Technical Review 12* (1979), 3–15.
6. B. Jesjesm, F. Kretz and H. Maitre, 'Statistical Study of Edges in TV Pictures', *IEEE Trans on Com. Vol. Com-27, No. 8* (Aug 1979), 1239–1247.

The following three colour plates are of a high resolution colour monitor displaying an electronically generated circular zone plate.

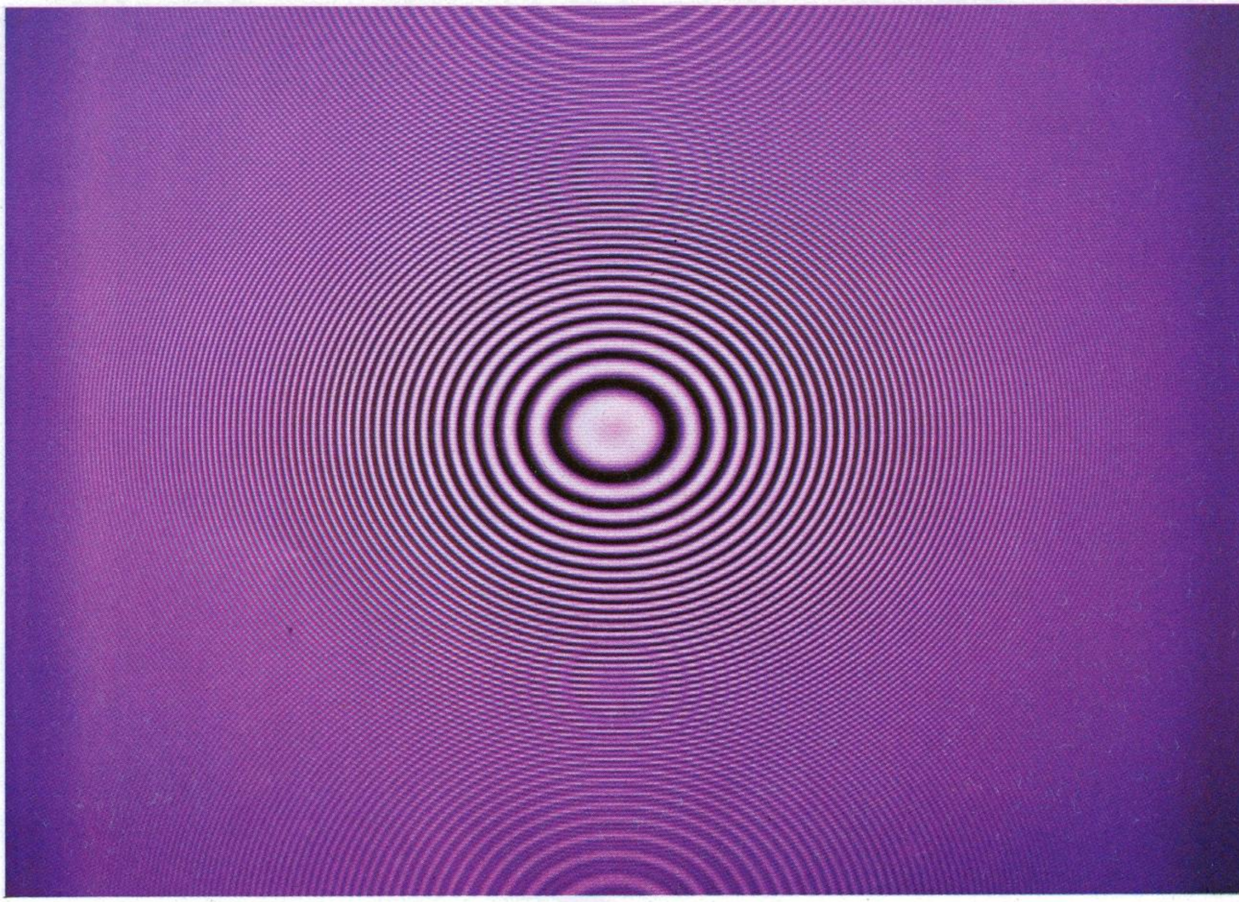


Plate 1. The original signal, slightly modified by horizontal filtering in the (14 : 14 : 14) MHz picture, is a continuous spatial frequency sweep which serves to reveal the effects of varying sampling frequencies and filtering in one and two dimensions.

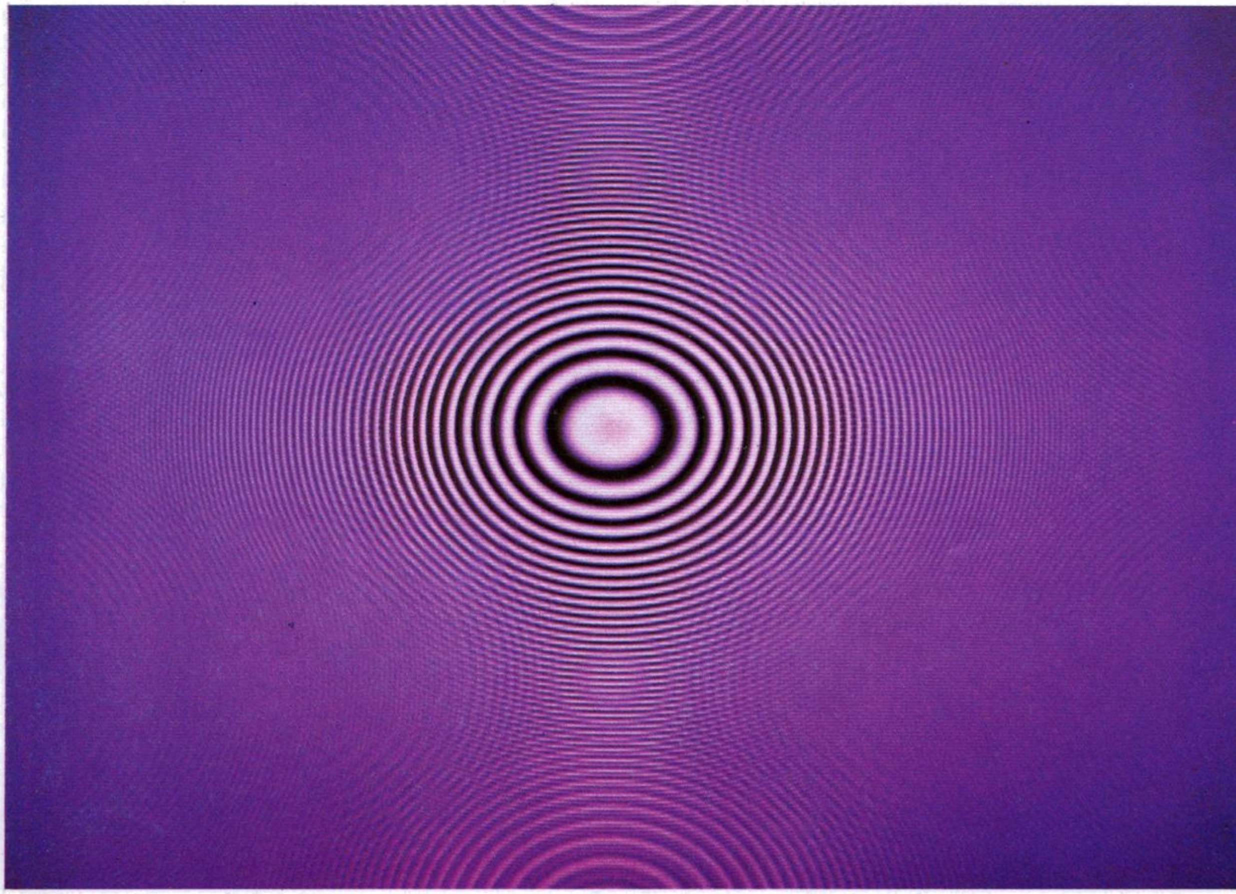


Plate 2. Effects produced by sampling frequencies of (7 : 3.5 : 3.5) MHz.

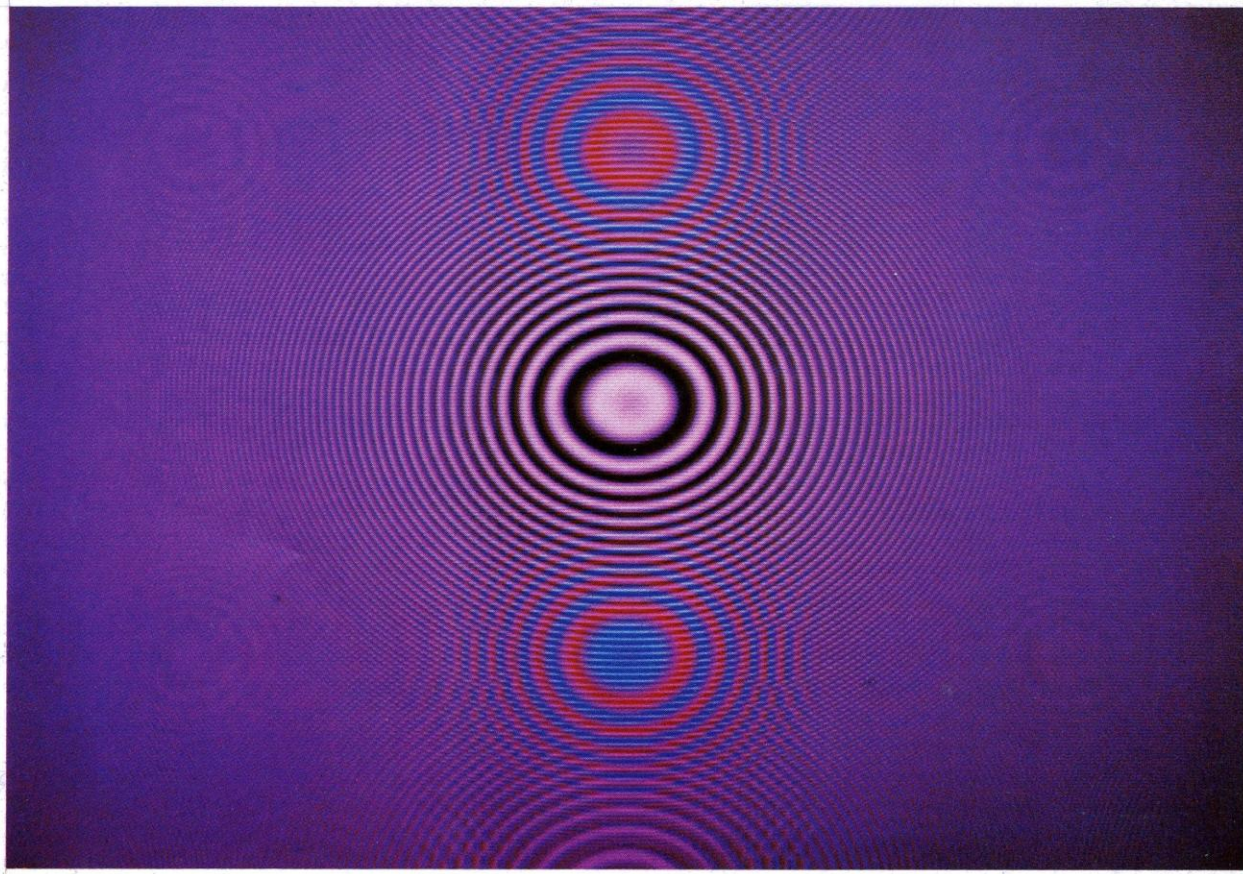


Plate 3. Effects produced by sampling frequencies of (9 : 3 : 0) MHz.

MIKE TOOMS, Dip.E.E., C.Eng., MIEE, MIERE, commenced his career with EMI in 1959. Subsequently he was with ABC Television, initially in technical operations and subsequently in system design and planning; with the IBA as Senior Quality Control Engineer sharing responsibility for the establishment and maintenance of technical codes of practice; with Rediffusion Television in Hong Kong as Controller of Engineering and Production Services, where in addition to the day-to-day operation he was responsible for colourising the studio and establishing a dual service transmitter network; and latterly, again with EMI, as Manager Broadcast Systems.

In 1977 he formed his own consultancy company Protel Broadcast Services Limited, since when, in addition to the usual consultancy activities, he has been deeply involved in the consideration of the systems problems associated with implementing an all digital studio centre.

Systems Engineering Considerations in the All Digital Television Production and Transmission Centre

by M. S. Tooms

Synopsis

A design study has been undertaken for an all digital television centre. This was considered the most practical method of exploring: (*a*) whether system configurations alternative to those adopted for analogue environments would be advantageous; (*b*) of evaluating what would be the consequences of these alternative configurations on the specifications of new digital equipment; (*c*) of determining how the phasing of the introduction of the digital system into a television operation could best be achieved, and (*d*) of assessing the suitability of a particular set of digital coding parameters for use in an overall system context.

This chapter emphasises those aspects of the study which pertain to the system configuration and the resulting desirable features of the digital equipment, particularly as they relate to the adoption of standards for multiplexing the video, the audio and the pulses within the installation.

In television terms, systems engineering embraces the design and implementation of the electronics system which supports the production processes. However, in order to discuss systems engineering in a meaningful manner in a field which is as yet virtually undefined, it is first necessary to outline in general terms an operational specification for the electronic system. It is likewise necessary to define the features of the digital techniques to be used in designing the system. It is desirable that the operational specification be sufficiently extensive to enable full exploration of any shortcomings in the proposed digital parameters.

Furthermore, it seems likely that, as the system configuration evolves, additional important features desirable in a set of digital parameters may be

highlighted. What we have, therefore, is a problem which requires an iterative solution. If we wish to minimise the possibility of constraints in future systems design it would be expedient to carry out a design exercise for a digital studio system before any proposed digital parameters are standardised.

This chapter describes the result of implementing one loop in this iterative process. The purpose at that stage was not to advocate a particular set of digital parameters, but to explore some of the desirable features which would be incorporated in any proposed set of parameters and in any digital equipment yet to be developed. Unless the system requirements are made known to the manufacturers during the specification stage, digital equipment might be introduced which would be the direct digital counterpart of current analogue equipment and which could, in consequence, place severe constraints on the flexibility of design of all future digital television centres.

DIGITAL PARAMETERS

Before the system design can proceed, a likely set of digital parameters must be selected on which to base the design.

Video

With regard to video, several sets have been proposed, and some have already been found inadequate.

Though initially the composite (NTSC, PAL) signal was favoured for coding, the advantages of component (Y, R–Y and B–Y) coding have become more clear, and this alternative seems likely to gain general acceptance. There also seems to be general acceptance that eight digits are required to satisfactorily define the level of each sample (at least in the area of the system beyond signal source processing). Both these features are assumed in the choice of a set of digital parameters for the purpose of this chapter. As to sampling frequency, one proposal which has been made, and which seems especially promising, is based upon a video sampling frequency of precisely 13.5 MHz.

This frequency seems to satisfy most criteria. It has the additional important property that it may be used for both 525-line and 625-line systems giving an orthogonal sampling structure and an identical number of 704 active samples per line in both systems. This fortuitous situation results from the effects of the slightly different line and line blanking periods cancelling out between the two systems. A more detailed review of some of the properties of this sampling frequency is contained in the Appendix.

Though significant differences between the systems will necessarily remain, a common sampling rate and a common number of active samples per line will each do much to simplify multi-standard equipment, which could eventually lead to significant reductions in cost.

As previously noted, eight bits are used to describe each video sample, leading to a nominal resolution of 1 part in 256. In fact, the luminance signal extends from black, equal to a digital level of 16, to white, equal to a level of 240. The zero levels of the colour difference signals are both given the values of 128, and the amplitudes are adjusted so that each peaks to the same minimum and maximum levels as the luminance signal, that is 16 and 240. The B–Y and R–Y signals suffer an attenuation with respect to the luminance signal of 0.562 and 0.714 respectively in order to achieve this result. The excursions of the signals in the digital domain are illustrated in Fig. 1.

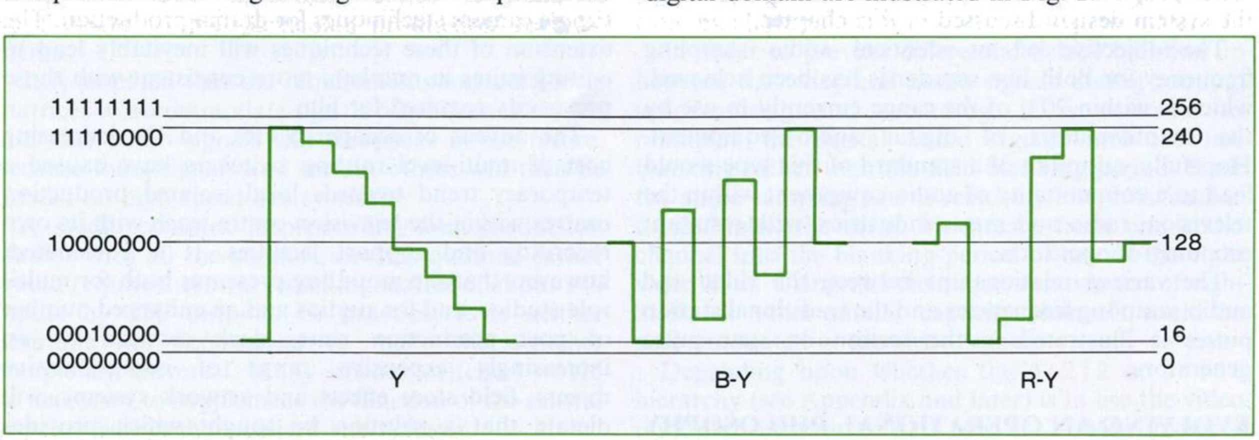


Fig. 1. Component signals in the digital domain.

The luminance and the two colour difference signals are time multiplexed on a bit basis leading to one video signal carried on eight wires. The sampling rate for the colour difference signals is discussed later.

The digital coding method is linear PCM with positive binary for luminance and offset binary for colour difference. The form of signal is non-return to zero.

These signals together with the various clock frequencies are likely to be carried around the system local to the central apparatus room via 25-way Canon D connectors and ribbon twisted pair cable.

On the basis of the foregoing and the further advantages which are highlighted in the forthcoming paragraphs on audio, these parameters have been adopted for evolving the system design described in this chapter.

Audio

With regard to audio we seem to be rapidly approaching the point of standardisation, though as far as is known, the requirements of digital audio in a television system environment have not yet been fully explored.

There appears to be a consensus for 16 digits to define the required number of audio levels. With regard to the sampling frequency, however, there seems to be a divergence of views as to whether this should be locked to the video system or be asynchronous.

Some of the factors which appear relevant to the selection of a digital audio sampling frequency are discussed in the Appendix. On the basis of this discussion a video locked sampling frequency of 60 kHz has been somewhat arbitrarily selected from those proposed in the Appendix for the purposes of the system design discussed in this chapter.

The objective of an identical audio sampling frequency for both line standards has been achieved, which is within 20% of the range currently in use by the manufacturers of digital audio equipment. Hopefully, adoption of a standard of this type would lead to a commonality of audio equipment within the television, radio and music industries, with resultant economy of operation.

The various relationships between the video and audio sampling frequencies and the traditional station pulses is illustrated in the section on sync pulse generators.

EVOLVING AN OPERATIONAL PHILOSOPHY

In order to explore the gamut of production requirements, a comprehensive television operation has been assumed which incorporates a number of studios in which drama, series, light entertainment, current affairs and news production may be mounted. In addition, both fully 'in house' post production facilities and a continuity suite for continuously feeding the local transmitter and the network are included.

As technology advances, so the requirements of production keep pace, always aiming at using the new technology to the limit in order to improve in one way or another the final programme. In this changing situation it is difficult to define precisely what form the future television production system will take. At best, current trends may be extrapolated to indicate the requirements of the future.

In the production of major programmes, for example, these trends indicate an increasing use of multi-track audio recorders both in the shooting and in the post production stages.

Though the cost of basic equipment such as cameras and VTRs has dropped in real terms over the years, use of such equipment has increased in a manner which has generally led to greater capital investment. Furthermore, recently developed equipment such as frame store effects, still stores and electronic artwork systems, represent still further capital investment.

The availability of ever more sophisticated post production systems is accelerating the trend towards selecting the system of shooting which matches the type of production. In general terms this means adopting single recording (multi-camera) techniques for live shows and economy series; multiple record (multi-camera) techniques for polished series and certain audience comedy shows, and sequential record (single camera) techniques for drama production. The extension of these techniques will inevitably lead to editing suites in numbers more consistent with those previously required for film.

The advent of cheaper VTRs and the increasing cost of multi-level routing switchers have caused a temporary trend towards local isolated production centres within the television centre, each with its own recording and support facilities. It is considered, however, that the mounting pressures both for multi-role studios, and for studios and an enhanced number of post production suites each serviced by an increasingly expensive range of comprehensive mixers, field store effects and artwork systems, will dictate that a solution be sought which provides greater flexibility and economy in operation.

These constraints infer a high utilisation of resources and the ability to direct them, as required, to the particular area of production which needs them at any given time.

OUTLINE SYSTEM CONFIGURATION

If the philosophy outlined in the previous paragraphs is accepted as the basis for the system design, the main features of the system configuration follow fairly naturally.

A Central Switching Matrix

Perhaps the primary requirement is to accommodate an increasing number of sources in an increasingly flexible manner. The obvious way of achieving this is to provide a central routing switcher which, in broad terms, enables any primary or secondary source to be selected to any destination. In order to accommodate stereo and separate music and effects tracks it is assumed that each source or signal in this arrangement will consist of a video signal plus four associated audio signals. Potentially, a number of very attractive features result from this configuration, a discussion of which is beyond the scope of this chapter.

Once a universal switcher of this type (together with the flexibility of switching arrangements now available) is tentatively accepted it becomes realistic to evaluate the possibility of also incorporating the normal pre-selection matrices of the studio mixers into the central matrix. This philosophy may be extended still further by arranging for the various sub-units of the mixers, i.e., the mid-effects units, the downstream key units and the frame store effects units, to be assigned to the control panels in the studio or post production control suites in numbers appropriate to the particular production requirements.

It is assumed that the requirements for multi-level matrices to accommodate signals other than the video and four audio signals will disappear in the future, because these signals and certain others will then be carried on high speed bus systems.

A studio complex incorporating four production studios with all the support services outlined earlier would require a central switching matrix of about 140 inputs by 250 outputs. In practice, it is neither desirable nor economical to make the switcher completely universal. Many smaller switchers would be necessary to supplement the function of the central switcher within the localised sub-systems.

A schematic of the conceptually simple configuration which results from these proposals is shown in Fig. 2.

The Audio System

The days when audio could be regarded by systems designers as a necessary adjunct to video are rapidly coming to an end. To fully service the multi-track requirement of the future, a multi-track system based upon an audio routing matrix which routes multi-track signals between the studios, dubbing theatres and recorders will be essential. A 24-track system is assumed for the purpose of this chapter. The extent of a system of this type is illustrated in the schematic of Fig. 3.

DIGITAL ASPECTS OF THE DESIGN

Implementation of the routing switchers described, if in analogue form, would be prohibitively expensive for the majority of users. In digital form, however, various techniques are easier to apply which make the proposals more feasible.

Video and Audio Multiplexing

In a complex environment of the type described, the single most useful move which could be made to reduce complexity and costs would be to multiplex together the video and four audio signals. By this means reducing, for example, a five level switcher carrying a video signal and four related audio signals, to a single level video plus four audio switchers. This approach has advantages also in routing married signals around the system.

The availability of cheap 8-gate devices dictates that the audio signals (and any other data that pertains to the signal) must be carried on the same eight wires as the video if costs are to be prevented from escalating.

In order to prevent differential delays building up between the video and audio signals during routing, synchronising and recording, it is desirable to multiplex the digital audio signals into the line blanking rather than the field blanking period. Since the audio sampling rate is relatively low, it is feasible to multiplex the serial bit stream from one audio channel into the blanking period of one of the eight video lines. Using this technique a maximum of eight audio or other similar capacity channels could be accommodated in the video signal.

Depending upon whether the 4 : 2 : 2 or 3 : 1 : 1 hierarchy (see Appendix and later) is in use the video multiplex bit-rate is 22.5 Mb or 27 Mb, giving a number of unused samples in the blanking period in

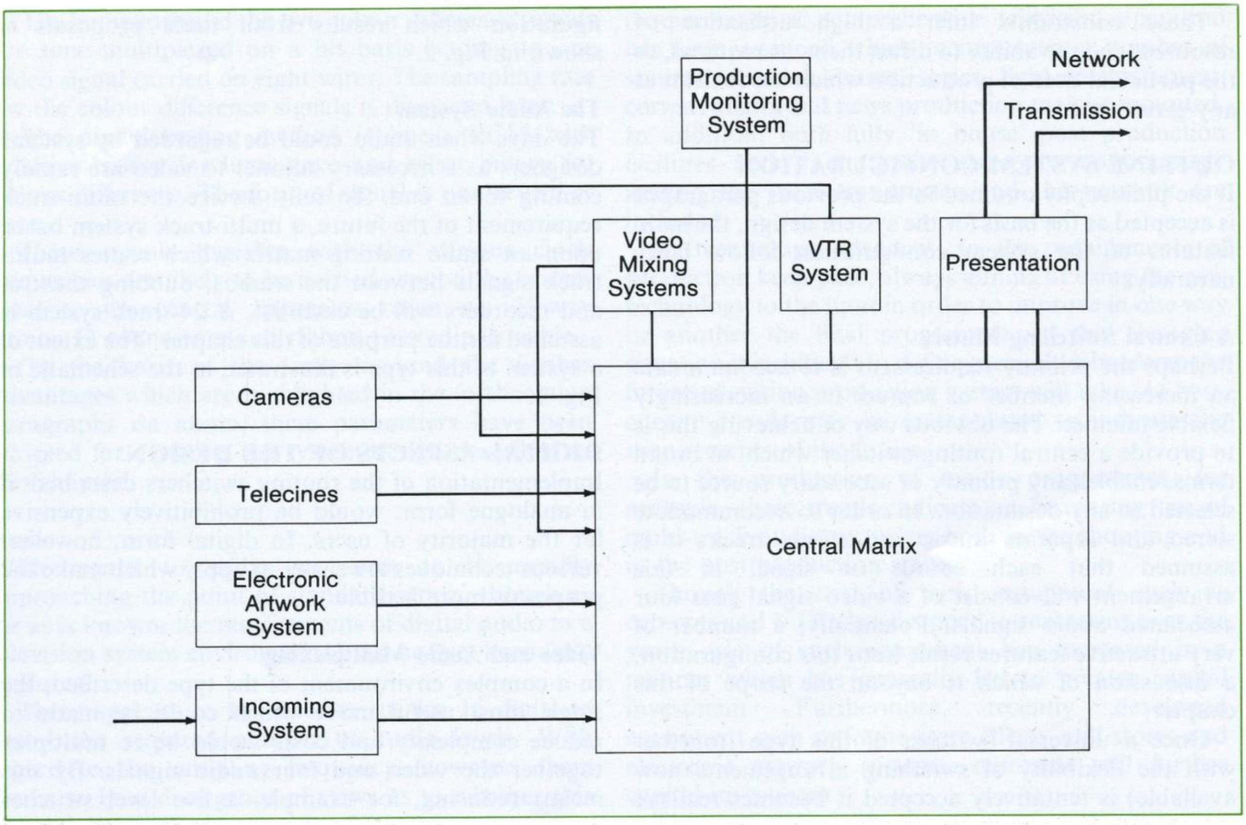


Fig. 2. Outline schematic of system configuration.

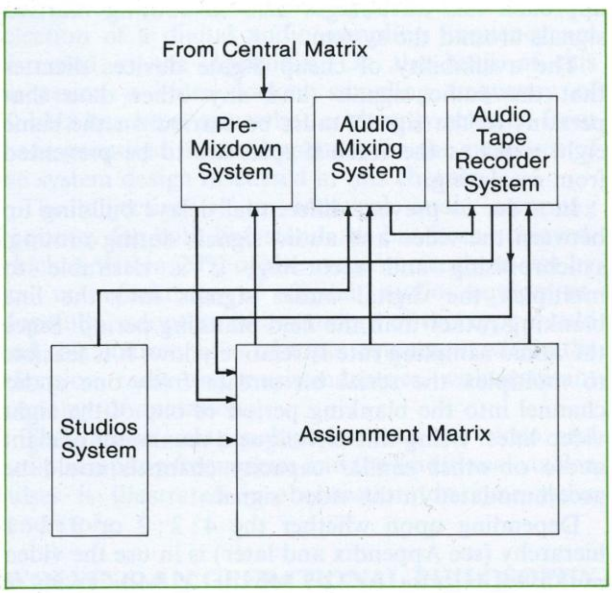


Fig. 3. Outline schematic of 24 channel audio system.

accordance with Table 1.

Each audio channel is sampled at 60 kHz which requires four samples per line and leads to the sampling patterns on the two-line standards illustrated in Table 2 of the Appendix. If the audio samples are serialised there will be 4×16, that is, 64 samples to be accommodated in each line blanking period. Thus, it can be seen that the audio uses only between 20% and 25% of the unused samples in each line, leaving plenty of capacity for line numbering, synchronisation etc.

The technique of multiplexing may be extended to the multi-channel audio system. Twenty-four channels of digital audio sampled at 60 Hz, and with 16 bits to define the level of each sample, will require an overall bit-rate of: $24 \times 60{,}000 \times 16 = 23$ Mbit/s.

Thus, the bit-rate per wire is in the same range as the multiplexed video plus four audio system, enabling use of the same technology for both types of routing switcher.

A Multiplex Standard?

It is very convenient in a medium-to-large sized

installation to convey the signals around the entire system in one of the two formats discussed: that is, video plus four audio multiplexed and 24 channel audio multiplexed, for the married video plus audio and audio only sub-systems respectively.

Since multiplexing needs be applied only at source, and de-multiplexing at station and monitoring outputs, the overall system would be greatly simplified by the adoption of a combined digital video and audio multiplex standard for use in the studio. It is envisaged that manufacturers of tape recorders, routing switchers, synchronisers and mixers would design their equipment, at least optionally, to accept and relay the signals in multiplex form. This would likewise supply to those equipments which might be needed to process the video and audio signals in accordance with different instructions, as, for example, in the presentation or continuity switcher.

Improving the Efficiency of the Switcher

A further technique which may be used to reduce the cost of the switcher is Clos's method[1] of dividing up a large switching matrix of crosspoints into a number of smaller matrices arranged in three columns, as shown in Fig. 4, where all matrices in the same column are of the same type. The central column acts as an exchange between the input and output matrices, which are located in the first and third columns respectively. It can be shown that, provided certain criteria are met, no input path is ever blocked to an output path.

Figure 4 shows how a system using these techniques may be configured for a 140×250 matrix. The centrally located figures indicate the number of individual matrices in each column, and the figures adjacent to each matrix indicate the numbers of inputs and outputs respectively. As can be seen, the total number of crosspoints is 14,060 compared with 35,000 for a traditional matrix. Naturally, the ideal numbers illustrated may require modification to accommodate any particular manufacturer's board configuration; nevertheless, a matrix of this size will generally bring about a saving of about 55% in crosspoints as compared with the traditional approach. Larger matrices lead to greater efficiencies.

Though this technique could be used, and in fact is increasingly being used for analogue signals, the need for the signal to traverse three sets of crosspoints, and any interfacing between groups of crosspoints, places a heavy demand on the performance of each section of the matrix. In digital terms, of course, the crosstalk and distortion suffered by the signal are irrelevant so long as these are below a relatively high critical level.

Chrominance Bandwidth

In order to provide a subjectively matched spatial resolution between the display of the luminance and chrominance components of a picture, the chrominance bandwidth need not exceed one-third of the luminance bandwidth. At first sight this would

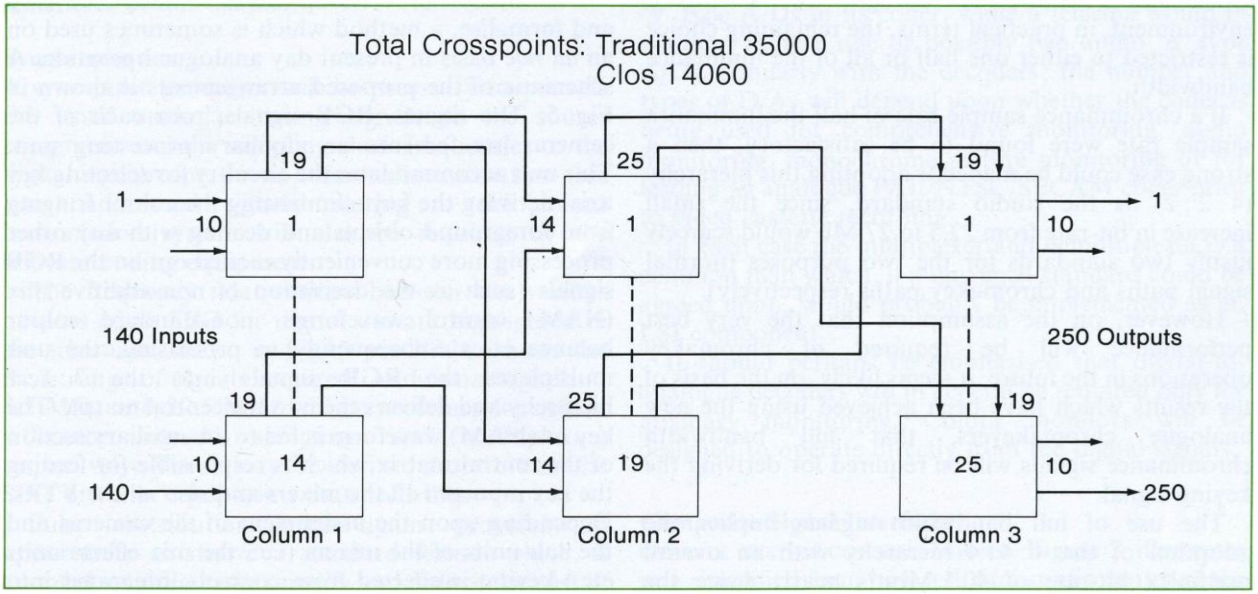


Fig. 4. Clos type arrangement for a 140×250 switching matrix.

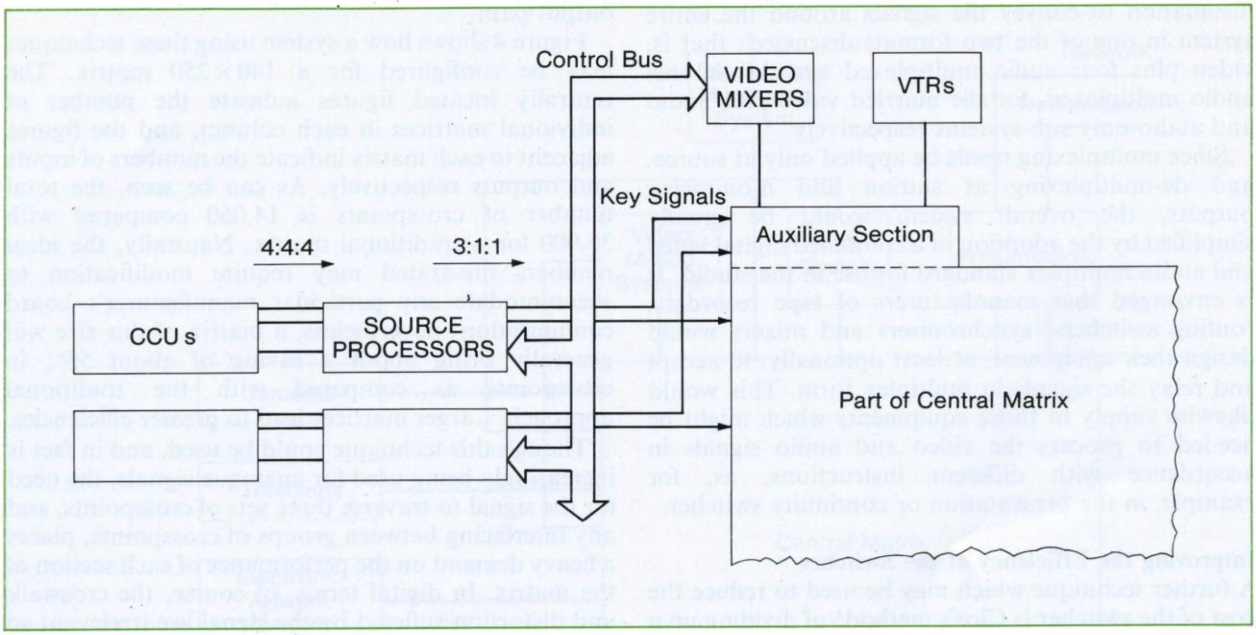


Fig. 5. Proposed chromakey arrangement using full chrominance bandwidth.

seem a reasonable figure to adopt in deriving the sampling frequency of the chrominance (or colour difference) signals in a digital system.

It has been shown, however[2], that for good chromakey, this bandwidth is inadequate. Work is progressing towards establishing the minimum satisfactory bandwidth; though, in a digital environment, in practical terms, the remaining choice is restricted to either one half or all of the luminance bandwidth.

If a chrominance sample rate of half the luminance sample rate were found to be satisfactory, then a strong case could be made for adopting this hierarchy (4 : 2 : 2) as the studio standard, since the small increase in bit-rate from 22.5 to 27 Mb would scarcely justify two standards for the two purposes (normal signal paths and chromakey paths respectively).

However, on the assumption that the very best performance will be required of chromakey operations in the future, it seems likely, on the basis of the results which have been achieved using the new analogue chromakeyers, that full bandwidth chrominance signals will be required for deriving the keying signal.

The use of full bandwidth signals implies the adoption of the 4 : 4 : 4 hierarchy with an overall multiplex bit-rate of 40.5 Mbit/s nearly twice the 22.5 Mbit/s rate of the 3 : 1 : 1 hierarchy system. Though this system would be ideal, it would of course impose tremendous burdens on the technology, inevitably delaying the introduction of an all digital studio centre and encouraging the production of 'non-standard' equipments for temporary solutions, to the detriment of the long-term situation.

One method of avoiding this is to adopt, develop and formalise, a method which is sometimes used on an *ad hoc* basis in present day analogue operation. A schematic of the proposed arrangements is shown in Fig. 5. The digital, RGB signals from each of the cameras are fed into an auxiliary processing unit. This unit accommodates the circuitry for selecting key axis, deriving the key, eliminating the colour fringing from foreground objects and dealing with any other processing more conveniently carried out on the RGB signals (such as the derivation of non-additive mix (NAM) control waveforms, non-standard colour balance etc.). Subsequently to processing, the unit multiplexes the RGB signals into the 3 : 1 : 1 hierarchy and delivers them to the central matrix. The key or (NAM) waveform is fed to an auxiliary section of the central matrix which is responsible for feeding the key inputs of all the mixers and also all the VTRs. Depending upon the assignment of the cameras and the sub-units of the mixers (i.e., the mix effects units etc.) keying is effected from controls integrated into the video mixer panels. For 'roll back and insert'

chromakeys, or wholly post-production chromakeys, the key signal is recorded upon a second VTR which is replayed in synchronism and selected as required via the central matrix for full bandwidth chromakeying. 'Downstream' chromakeys are not normally as critical as 'in house' chromakeys and would use signals derived from the multiplex 3 : 1 : 1 signals. Nevertheless, compared with analogue downstream chromakeys, the results would still be significantly superior because of the wider chrominance bandwidth and the absence of crosstalk.

Though these proposals serve to complicate the control system this is an area in which microprocessors are ideally suited to the task. It is considered, therefore, that the simplification of the signal paths and reduction in the time scale of introduction of suitably advanced technology together justify the enhanced complexity and resulting flexibility of the proposals.

In summary it is advocated, all other factors being equal, that if it is found that satisfactory chromakey can always be attained by using half luminance bandwidth, then a 4 : 2 : 2 hierarchy be adopted for the studio centre. If, however, full bandwidth chrominance signals are required, then it is recommended that the more economical 3 : 1 : 1 hierarchy be adopted and chromakey be dealt with by supplementing the overall studio system configuration. For convenience the 3 : 1 : 1 hierarchy is selected completely arbitrarily for the purposes of the remainder of this chapter.

Synchronisation

A sync pulse generator will be necessary to provide the usual analogue timing signals to the source equipment and the sampling clock frequencies to the digital equipment. In an intermediate situation (see later) additional signals relating to the colour transmission system will also be required by the relevant coding equipment.

The frequency relationships between the various outputs of an SPG for digital use are highlighted in the schematic shown in Fig. 6 which also shows the basic range of signals required. Additional signals at 4.5 MHz and 22.5 MHz will be necessary for colour difference signals sampling and multiplexing respectively. These may be derived from either the SPG or other equipments that need them.

Pulse distribution once again provides an opportunity to derive a standard format which, if adopted, could greatly simplify and reduce the cost of future installations.

INTERFACING TO THE ANALOGUE DOMAIN

Component Signal Interfaces

Although technically feasible, it might be practically unrealistic to expect digital cameras to be available for installation in the early all digital studio centres[3]. Accommodating analogue cameras, however, causes no difficulty or compromise since, in system terms, these devices are situated on the periphery and do not react with the system.

It might also be expedient, during the intermediate period, to make provision for monitoring digital signals on analogue devices. The necessity to provide for these functions means that analogue-to-digital codecs will be required; at least, initially.

Analogue-to-Digital Coding and Decoding

If account is taken of the digital-to-analogue decoders (D/As) required for various monitoring purposes a large number of equipments for several different purposes are necessary. Hopefully, if the common features can be defined it might be possible to design a standard pair of units which may be supplemented to suit particular requirements.

Figure 7 illustrates the functions required of these coders (but not, it should be emphasised, the manner in which the functions may be realised). The number and types of A/Ds and D/As used will depend upon the purpose for which the coder is required. For example, a camera would need only three of the video 'V' type A/Ds in its coder, while a telecine would in addition require one or possibly two audio 'A' type A/Ds. Similarly with the decoders; the number and types of D/As will depend upon whether the coder is being used for comprehensive monitoring, audio monitoring, monochrome picture monitoring or for feeding an analogue PAL/NTSC/SECAM coder prior to video cassette recording.

In each case the multiplexer and demultiplexer is a common item which, if a universal standard could be agreed upon, would probably be implemented as an LSI component.

It should perhaps be emphasised that digital-to-analogue decoders will in general be required only for auxiliary monitoring. Colour monitors will be designed to operate directly from the digital signals.

Composite Signal Interfaces

OVERALL SYSTEM CONSIDERATIONS. Figure 2 illustrates in a simplified form how a digital studio centre of the future might appear. All equipment is digital; and it is

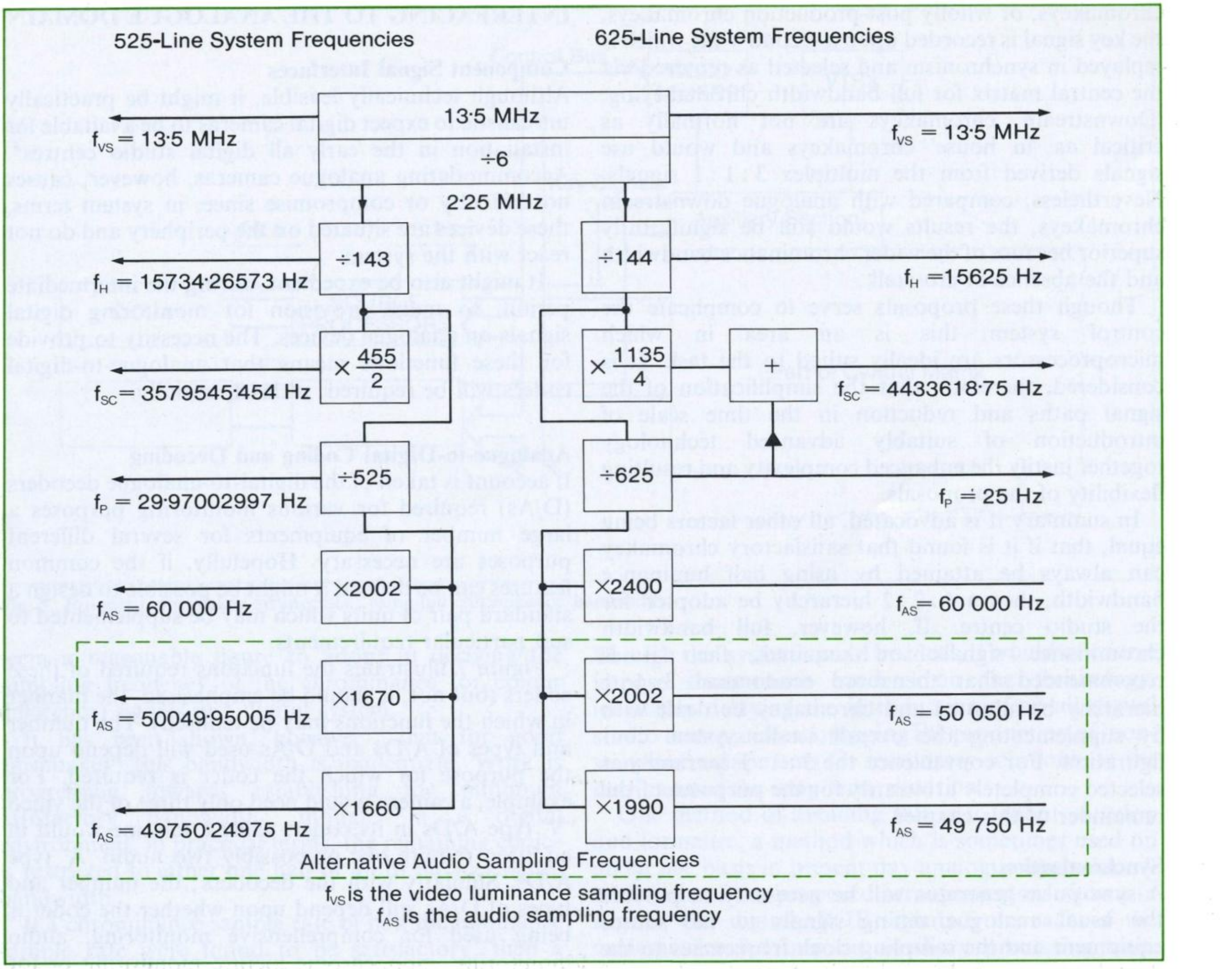


Fig. 6. Schematic of sync pulse generator for a digital studio system.

assumed that incoming signals are presented in the format defined for the purposes of this chapter. Similarly, outgoing signals will follow the same format, the transmission feed being digitally encoded to the composite PAL/NTSC format at the transmitter before undergoing digital-to-analogue conversion. In the case of SECAM the digital-to-analogue conversion would take place before coding to the SECAM format. (This is always true for SECAM wherever reference is made to PAL/NTSC coding later in the text.) Such a system is conceptually simple yet provides tremendous operational flexibility and consistently excellent quality of pictures and of sound. System timing would be greatly simplified by ensuring that all signals entering the system were made synchronous and by applying automatic delay correction at the inputs of each of the mixing systems.

Unfortunately, in practical terms, the change to this delightfully simple configuration necessitates a detour through the technologically complex and untidy intermediate stage where both composite and component signals are embraced in the same system.

There seems no way of avoiding this intermediate stage if the full advantages of digital are to be achieved within a reasonable time-scale.

Even if it is decided to effect a revolutionary change to digital by building a new system, the factors listed below will ensure that the vagaries of successive

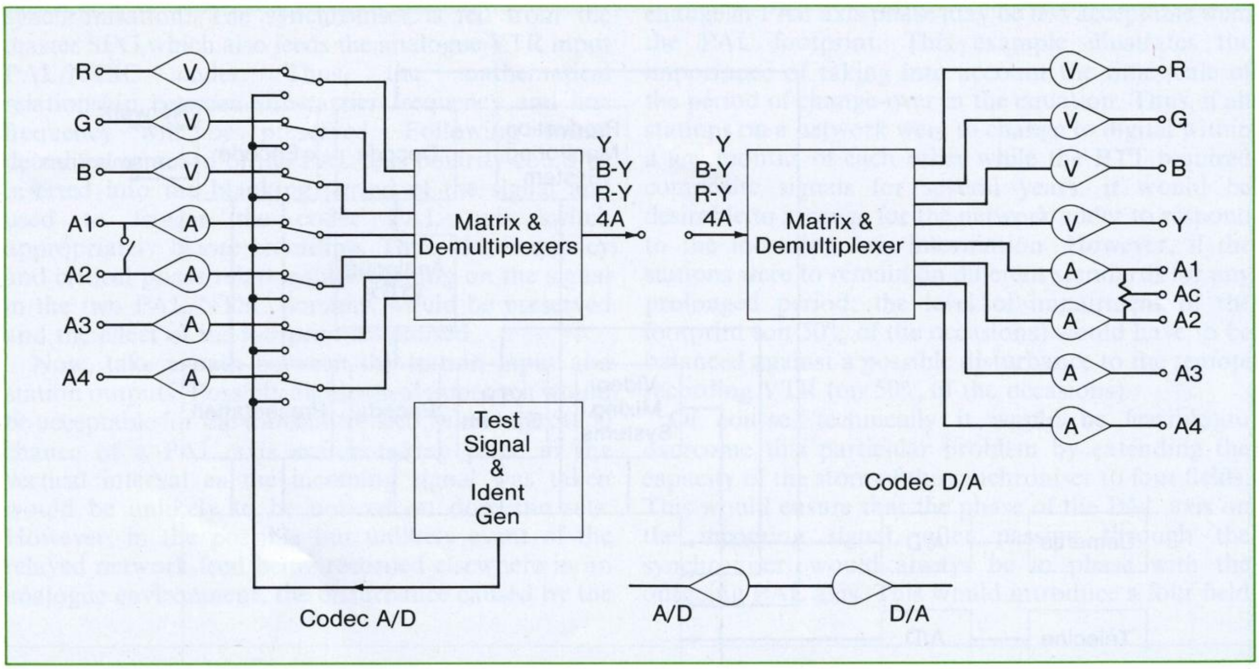


Fig. 7. Schematic of analogue-to-digital coders and decoders in a TV system.

transformations of the signals, from the component to the composite format, must be suitably accommodated in the design if the impairments associated with this procedure are not to offset the advantages in quality gained by going digital.

These factors are:

(i) the requirement to accept and offer signals to the PTT authority in a composite format, whether analogue or digital, possibly for many years to come.

(ii) the relaying and recording respectively of historic analogue composite tapes and composite tapes for distribution to non-digital organisations, and,

(iii) as we have seen, the accommodation of otherwise perfectly satisfactory analogue source equipment, such as cameras and telecines, in the digital environment.

These requirements modify and complicate the simple system as can be seen in Fig. 8. Digital PAL/NTSC decoders appear at the station inputs and analogue VTR outputs, and digital PAL/NTSC coders at the station outputs and analogue inputs.

PAL/NTSC/SECAM IN THE DIGITAL STUDIO CENTRE. Dealing with a composite signal, and particularly PAL in a luminance and colour difference environment, is perhaps the most demanding aspect of system design. Unless sufficient care is taken at the system design stage, serious impairments can occur.

The problem is that in composite systems the chrominance signal and the higher frequency luminance signals share interleaved spectrum space, and, for moving pictures, identical spectrum space; thereby making impossible the complete separation of the signals during demodulation. The result is that inevitably there will be crosstalk between the signals in both directions. Normally, this manifests itself as crosscolour and cross luminance; however, if the separated signals are re-coded back to PAL/NTSC, generally speaking, and particularly for PAL, the effect of the distortion is exacerbated. The effects of this type of crosstalk in a re-coded PAL signal are referred to as the PAL footprint. In a complex television environment such as is found in the UK, once the change to digital has been made in the studio, the signal may suffer many PAL coding and de-coding processes in tandem before being finally transmitted, and this could lead to a very significant cumulative impairment.

There are several ways in which the fundamental problems can be tackled:

(i) the use of 'clean PAL' techniques which are based upon combing luminance and chrominance

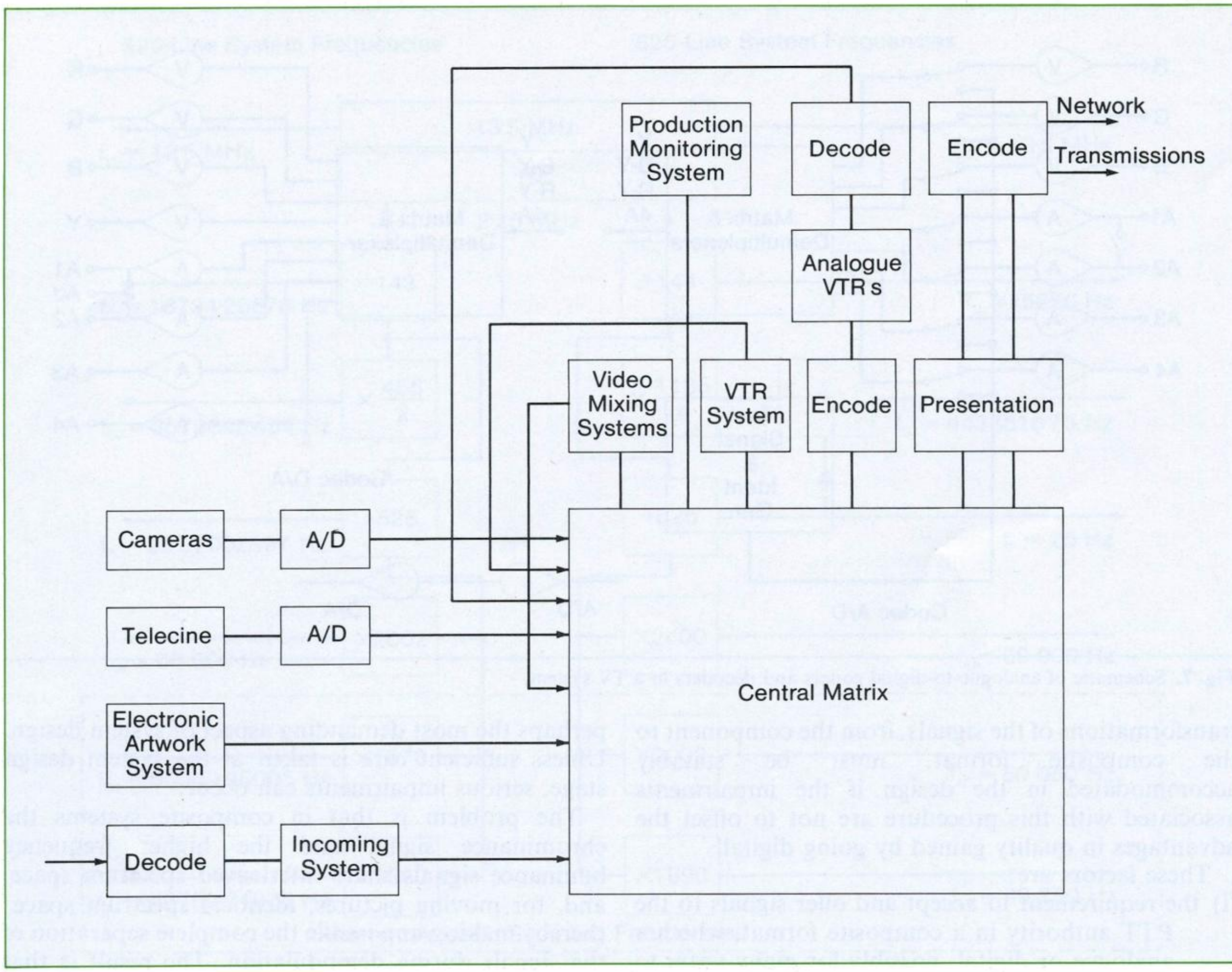


Fig. 8. Outline schematic of an intermediate system configuration.

signals before combination in the encoder, so that they occupy adjacent but not identical spectrum space.

(ii) the use of advanced decoding techniques which reduce the level of the footprint, and,

(iii) by taking care within the system to re-code the signals using the same related sub-carrier frequency and PAL axis phase as in the original signal, so minimising the effect of the footprint.

Unfortunately, all these techniques have imperfections which necessitate a compromise between the footprint impairments and the impairments introduced by the adoption of the technique.

The use of clean PAL techniques will, without doubt, ameliorate the situation; however, care must be taken in selecting a comb characteristic for the filters which produces the best compromise between footprint and loss of spatial and temporal resolution.

Care must also be taken when using adaptive decoding techniques to ensure that other aspects of the performance are not unduly compromised.

CARE IN SYSTEM DESIGN. Looking again at Fig. 8 it may be possible to select from the foregoing methods of improving the footprint a different and more satisfactory compromise for each of the different functions necessitating NTSC/PAL coding.

Take, for example, the path between the station input and an analogue videotape recorder. The composite signal is converted from analogue to digital form and PAL/NTSC decoded prior to

synchronisation. The synchroniser is fed from the master SPG which also feeds the analogue VTR input PAL/NTSC coder. Thus, the mathematical relationship between sub-carrier frequency and line frequency will be preserved. Following initial decoding, a record of the PAL axis polarity could be inserted into the blanking period of the signal and used to trigger the coder PAL axis switch appropriately before recording. Thus, the frequency and critical phase relationships existing on the signal in the two PAL/NTSC domains would be preserved and the effect of the footprint minimised.

Now, take a path between the station input and station outputs. Possibly an identical approach would be acceptable for the transmitter feed, where the 50/50 chance of a PAL axis switch taking place in the vertical interval as the incoming signal was taken would be unlikely to be noticed on domestic sets. However, in the possible but unlikely event of the relayed network feed being recorded elsewhere in an analogue environment, the disturbance caused by the change in PAL axis phase may be less acceptable than the PAL footprint. This example illustrates the importance of taking into account the time-scale of the period of change-over in the equation. Thus, if all stations on a network were to change to digital within a few months of each other while the PTT required composite signals for several years, it would be desirable to arrange for the network coder to respond to the incoming axis information. However, if the stations were to remain on different standards for any prolonged period, the level of impairment of the footprint (on 50% of the occasions) would have to be balanced against a possible disturbance to the remote recording VTR (on 50% of the occasions).

Of course, technically it would be feasible to overcome this particular problem by extending the capacity of the store of the synchroniser to four fields. This would ensure that the phase of the PAL axis on the incoming signal, after passing through the synchroniser, would always be in phase with the outgoing PAL axis. This would introduce a four field

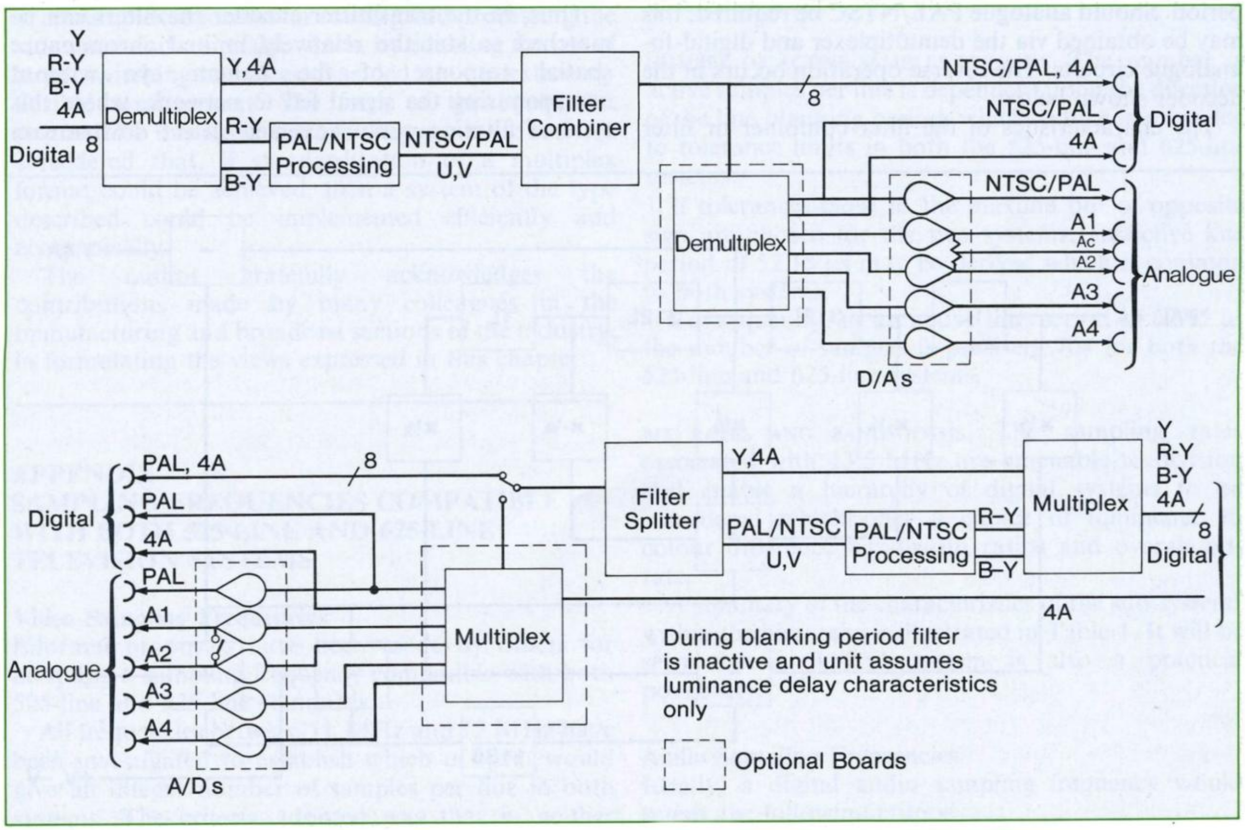


Fig. 9. Schematic of digital PAL/NTSC coders and decoders.

sequence into the digital domain, but, since only PAL signals arriving from remote sources would be affected, this would be unlikely to cause any operational compromise.

DIGITAL PAL/NTSC CODING AND DECODING. Figure 9 illustrates the functioning of a coder and decoder in the digital environment described. Since the PTT may require signals to be provided in either analogue composite form, digital PAL/NTSC or digital component format, these modular coders provide a universal solution. In the coder at the top of the diagram the multiplexed luminance, colour difference and audio signals are first demultiplexed into luminance plus audio, and colour difference signals are modulated onto the composite sub-carrier, and both sets of signals are shaped in a complementary fashion in the filter before being combined. The filter is rendered inactive during the line blanking period thus preserving the integrity of the audio and other signals. The filter output is a digital PAL/NTSC signal with audio multiplexed into the line blanking period. Should analogue PAL/NTSC be required, this may be obtained via the demultiplexer and digital-to-analogue circuits. The reverse operation occurs in the decoder shown below.

The characteristics of the filter combiner or filter splitter are defined to suit the task being undertaken by the equipment.

IMPROVING THE RECEIVED PICTURE. In using clean PAL/NTSC techniques the potential arises for not only eliminating the footprint but also significantly improving the picture as presented to the viewer, particularly those viewers who will invest in comb filter hi-fi sets which will become available.

Clean PAL/NTSC techniques can, of course, be fully exploited only between areas originating and terminating signals in the separate luminance and colour difference format. Frequently, when networking from an unmodified centre or taking material from analogue VTRs, only clean PAL/NTSC decoders will be in operation since the encoders are, or would have been, in an entirely analogue environment.

It is clear from the foregoing that, in order to achieve optimum separation of the luminance and chrominance signals, separation filters of different characteristics may be desirable for different purposes.

Thus, in the transmitter encoder the filters can be matched to suit the relatively limited chrominance spatial response of the human eye without compromising the signal fed to network, where this type of filtering could adversely affect downstream

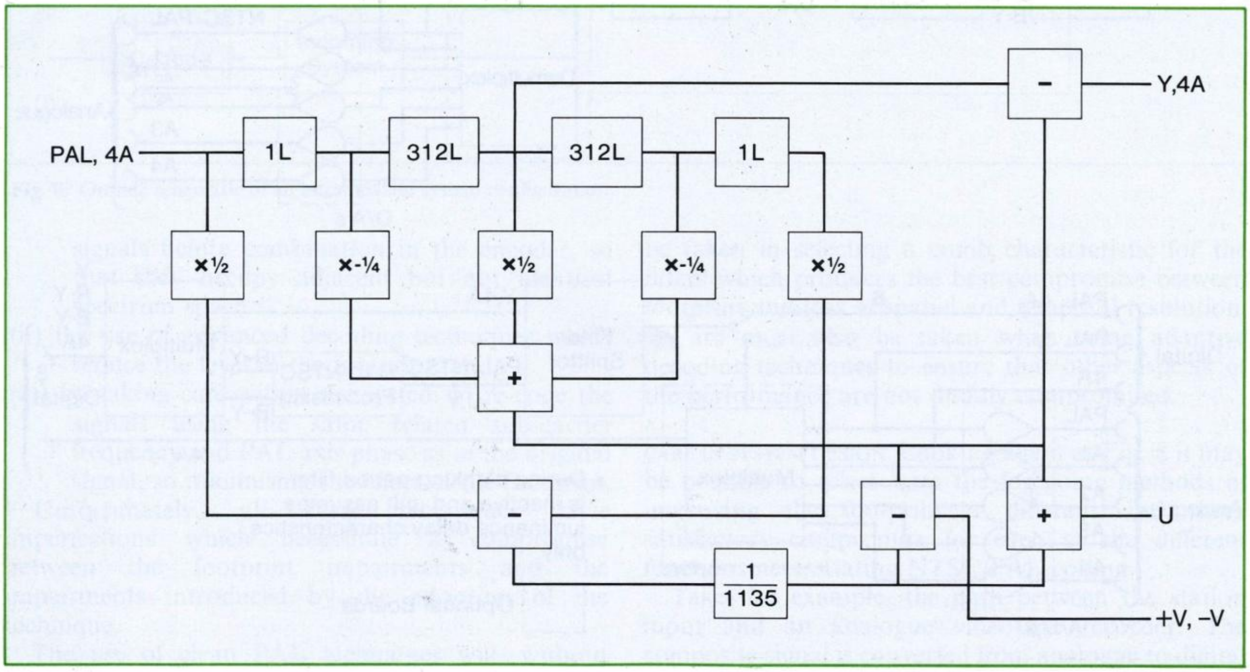


Fig. 10. Field delay component separation filter for PAL signals.

processing. The result on domestic sets could be a very considerable improvement in coarse cross colour performance and a virtual elimination of all cross colour from hi-fi sets.

Similarly, in the analogue VTR decoder, and possibly the coding equipment associated with external signals, a filter based upon picture delays would give almost perfect separation of the signals. A field delay version of such a filter giving a compromise between spatial and temporal response is shown in Fig. 10. Adaption techniques can be used to modify the response of these filters to match the picture content.

CONCLUSIONS

A practical system configuration has been evolved for an all digital television studio centre based upon the use of a set of coding parameters which satisfy the majority of the criteria hitherto specified and additionally provide a solution for both 525-line and 625-line systems.

The configuration is also based upon the adoption of a multiplexed format for routing and processing the video and audio signals; which, by enabling the use of large switching matrices, has led to an extremely flexible arrangement. The form of the equipment required for this approach has been outlined; and it is considered that, if standardisation of a multiplex format could be achieved, then a system of the type described could be implemented efficiently and economically.

The author gratefully acknowledges the contributions made by many colleagues in the manufacturing and broadcast sections of the industry, in formulating the views expressed in this chapter.

APPENDIX
SAMPLING FREQUENCIES COMPATIBLE WITH BOTH 525-LINE AND 625-LINE TELEVISION SYSTEMS

Video Sampling Frequencies

Informal proposals have been made by others for adopting a sampling frequency compatible with both 525-line and 625-line standards.

All frequencies between 11 MHz and 15 MHz have been investigated to establish which of them would give an integer number of samples per line in both systems. The criteria adopted was that in neither system should the specified tolerance of the resulting sub-carrier be exceeded in order to meet the integer relationship.

Only two frequencies were found to meet the above criteria, these were: 11.25 MHz precisely and 13.50 MHz precisely.

In fact, the relationships are exact and occur at every 2.25 MHz. The 11 Hz of tolerance available between the two sub-carriers is not utilised.

On the basis that 11.25 MHz is too low a sampling frequency for the 625-line system (particularly System I), the 13.50 MHz frequency has been adopted in this section in evolving a configuration for an all digital studio system.

NUMBER OF SAMPLES PER LINE. In the 625-line system: ${}^{f}h525 = 15{,}734.26573$ Hz (that is

$$\frac{4.5 \times 10^6 \text{ Hz}}{286}$$

and there are therefore

$$\frac{13.5 \times 10^6}{15{,}734.26573} = 858$$

luminance samples per line).

NUMBER OF ACTIVE SAMPLES PER LINE. The number of active samples per line is dependent upon the duration of the line blanking period, which in turn is subjected to tolerance limits in both the 525-line and 625-line systems.

If tolerances close to the maxima but of opposite sign are chosen for the two systems, an active line period of 52.15 μs may be derived which is common to both systems.

Consequently, in an active line period of 52.15 μs the number of samples is precisely 704 for both the 525-line and 625-line systems.

BIT RATES AND BANDWIDTHS. The sampling rates associated with 13.5 MHz are amenable to division and enable a hierarchy of digital systems to be considered which offer a choice of luminance to colour difference bandwidth ratios and overall bit-rate.

A summary of the characteristics of the sub-systems within the hierarchy is illustrated in Table 1. It will be seen that a 3 : 1 : 1 system is also a practical possibility.

Audio Sampling Frequencies

Ideally, a digital audio sampling frequency would satisfy the following criteria:

(i) The frequency would be high enough to allow

TABLE 1: BIT-RATES AND BANDWIDTHS OF SYSTEM HIERARCHY

HIERARCHY	SYSTEM	LUMINANCE SAMPLES PER LINE	LUMINANCE NOMINAL BAND-WIDTH	COLOUR SAMPLES PER LINE	COLOUR NOMINAL BAND-WIDTH	TOTAL SAMPLES PER LINE	BIT-RATE PER WIRE Mb	OVERALL BIT-RATE Mb
BASED UPON 13.5 MHz LUMINANCE SAMPLING								
4 : 4 : 4	525	858	6.0	858	6.0	2,574	40.50	324
	625	864	6.0	864	6.0	2,592	40.50	324
4 : 2 : 2	525	858	6.0	429	3.0	1,716	27.00	216
	625	864	6.0	432	3.0	1,728	27.00	216
4 : 1 : 1	525	858	6.0	214½	1.5	1,287	20.25	162
	625	864	6.0	216	1.5	1,296	20.25	162
3 : 1 : 1	525	858	6.0	286	2.0	1,430	22.50	180
	625	864	6.0	288	2.0	1,440	22.50	180
BASED UPON 6.75 MHz LUMINANCE SAMPLING								
2 : 1 : 1	525	429	3.0	214½	1.5	858	13.50	108
	625	432	3.0	216	1.5	864	13.50	108

adequate audio bandwidth without aliasing distortion.

(ii) The frequency would be equally acceptable to the television, radio and music industries.

(iii) There would be an integer number of samples per television frame in order to simplify multiplexing and editing procedures.

(iv) The frequency would give an integer number of samples per frame in both 525-line and 625-line systems.

(v) The number of samples per frame should be divisible by factors of the number of lines per frame to further simplify the multiplexing.

(vi) The frequency would not be so high as to be wasteful of bit-rate capacity.

Though it may not be possible to satisfy all these criteria, an approach as close as possible is obviously desirable.

CONSIDERATION OF AUDIO SAMPLING FREQUENCY CRITERIA. The nominal bandwidth of audio s[illegible]ls within a television operation is 15 kHz. On this basis an audio sampling frequency of about 32 kHz or twice line rate would seem to be about right. However, the professional audio industry, who are beginning to use digital equipment, have found the need to leave themselves plenty of headroom, and frequencies within the range of 44–51 kHz have so far been adopted.

All frequencies between 30 kHz and 62 kHz were investigated in terms of their ability to produce an integer number of samples per frame in both the 525-line and 625-line systems. Only the frequencies of precisely 30 kHz and 60 kHz were found to offer exact integers in both systems. However, a number of other frequencies were found to have values for the two line standards differing by less than one hertz. These frequencies fall into groups of thirteen centred at every 5 kHz, or more precisely at every 4,975 Hz, between 30,000 Hz and 44,925 Hz and between 45,075 Hz and 60,000 Hz. The frequencies within each group are separated by nearly 150 Hz. The characteristics of the

TABLE 2: AUDIO SAMPLING FREQUENCIES

525 LINES						625 LINES						
fa Hz	SAMPLES PER FRAME	Nr	Na	Nb	Nc	fa Hz	SAMPLES PER FRAME	Nr	Na	Nb	Nc	DIFFERENCE FREQUENCY Hz
30,000.00000	1,001	7		N/A		30,000	1,200	25	24	—	—	—
34,975.02497	1,167	3	155	1	19	34,975	1,399	1	476	149	—	0.02497
39,950.04995	1,333	1	242	283	—	39,950	1,598	1	277	348	—	0.04995
44,925.07493	1,499	1	76	449	—	44,925	1,797	1	78	547	—	0.07493
45,074.92907	1,504	1	71	454	—	45,075	1,803	1	72	553	—	0.07493
50,049.95005	1,670	5	41	84	—	50,050	2,002	1	—	498	127	0.04995
55,024.97502	1,836	3	—	88	87	55,025	2,201	1	—	299	326	0.02498
60,000.00000	2,002	7	—	14	61	60,000	2,400	25	—	4	21	—
49,750.24973	1,660	5	—	88	17	49,750	1,990	5	—	102	23	0.24973
48,011.98801	1,602	3	1	164	10	48,000	1,920	5	—	116	9	11.98801

$$fa = Nr\,(2Na+3Nb+4Nc)\,fp.$$

where fp = picture frequency

centre frequency in each group, together with one or two other interesting frequencies, are illustrated in Table 2.

SELECTING AN AUDIO SAMPLING FREQUENCY. The choice would seem to be between identical frequencies on both 525-line and 625-line systems; which, unfortunately, are either much below or just above the range of current practice, or a number of frequencies which are marginally non-identical but fall within this range.

Let us, somewhat arbitrarily, narrow the choice and consider in more detail the ramifications of adopting one or other particular frequency. Of the exact frequencies 30 kHz would give an upper audio frequency of about 14 kHz; which though marginally satisfactory for television, would be unacceptable for the music industry. Sixty kHz is some 20% higher in frequency than the range of current practice and would be somewhat wasteful of bit-rate, though otherwise it is satisfactory.

Of the non-identical frequencies, unless some hitherto unforeseen factor plays a part, the choice is based upon a frequency with a minimum divergence between the line standards and situated in the middle of the range of current practice. Referring to Table 2, the frequencies at nominally 50,050 Hz and 49,750 Hz meet this criteria.

The frequencies at 49,750 Hz have a divergence of about 0.25 Hz and provide a number of samples per field which are easily factorable. On the other hand, the frequencies at 50,050 Hz have a divergence of only 0.05 Hz, i.e. one part in a million; but the number of samples per field is not factorable on the 625-line system. Both, therefore, have attractions.

The choice then is between 60,000 Hz and two different frequencies at about 50,000 Hz.

It would seem essential, if digital audio mixers, processors and recorders are to be made available as economically as possible, that a common sampling frequency be adopted by all users. If the music industry could be persuaded to adopt 60,000 Hz on the basis that the 20% additional bit-rate is not critical in terms of technology or cost, then this would seem the ideal choice.

However, non-identical frequencies which differ by only 0.05 or 0.25 Hz are unlikely to be any less useful in the large majority of operations. In any local or national environment, equipment would operate totally satisfactorily using either of the two frequencies at about 50,000 Hz. Only when there occurred an international interchange which embraced a standards conversion would the 0.05 or 0.25 Hz become noticeable. Even in this case, the difference is unlikely to be a problem, since the

average delay through the standards converter exceeds the build-up in timing error caused by the 0.25 Hz difference in frequency for programmes of average length. The timing errors have values of 3.6 ms/hour and 18 ms/hour respectively.

The frequency of 48,000 Hz has also received attention as a candidate for standardisation. On 625-line systems it produces an integer number of samples per frame. However, on 525-line systems, the use of 29.97 Hz as opposed to 30 Hz for the picture frequency leads to a sampling frequency of 48,011.98801 if the integer relationship is to be retained.

This difference of 12 Hz is again unlikely to cause a problem in any local or national environment or indeed in the conversion of videotapes. However, in live international exchanges audio interpolators would be required to ensure no timing errors were introduced. This frequency is the only one (with the exception of 60,000 Hz) which also produces an integer number of samples every 1/24th of a second.

On this basis, 60,000 Hz and nominally 50,000 Hz and 48,000 Hz could all be accommodated. A final choice must be made on the basis of both the relative importance of the factors indicated in the preceding paragraphs and any other factors which have not been considered here.

References

1. C. Clos, 'A Study of Non-Blocking Switching Networks', *Bell System Tech. J.*, **32**, 406–424 (March 1953).

2. R. Rawlings, 'Chromakey in a Digital System', *International Broadcast Engineer, Vol. II*, **63**, 30–33 (September 1980).

3. M. S. Tooms, 'Going Digital', *Royal Television Society Journal, Vol. 18*, **4**, 5–9 (July/August 1980).

G. M. DRURY,
BA, B.Eng., C.Eng., MIEE, was born in Wales and graduated in 1967 from Liverpool University. He then joined a telecommunications research department at GEC Hirst Research Centre working on high-speed digital transmission systems. In 1972 he joined the Post Office Space Systems Development Department to study applications of high-speed digital techniques to space satellite systems. In 1974 he joined the Video and Colour Section of the IBA at Crawley Court where he began investigating the transmission of digital broadcast signals. In 1978 he became responsible for the digital video recording project team and continues to work in this area. He has also continued the study of transmission problems and, since 1978, has been a member of EBU specialist study group V1/Trans. In 1980 he successfully completed a BA degree with the Open University, having studied general arts subjects. He is married and has two daughters, whose attentions leave little time for his other interests of reading and music.

Digital Videotape Recorders for Component Coded Signals

by G. M. Drury

Synopsis
The videotape recorder has become a vital element in modern television programme production. It has evolved to fulfil various roles reflecting numerous aspects of the production process. The prospect that digital technology might, perhaps within the near future, begin to dominate broadcasting, naturally brings into question the practical feasibility of digital videotape recorders.

This chapter reviews the work done during the past few years to investigate digital video recording. This supports other work on digital video coding standards, to assist in identifying the constraints, set by recording technology, on future digital studio systems. This work is not yet complete, but considerable headway has been made in clarifying key design parameters, and in developing certain necessary techniques. These signal processing techniques and the problems which they solve, are here discussed, together with some of the theoretical factors of digital recorder design.

Introduction

In recent years considerable effort has been applied to the investigating of digital videotape recording. The techniques developed have led to some successful demonstrations employing experimental equipment. This equipment has been designed around existing tape transport mechanisms conforming to the A, B and C format codes of the Society of Motion Pictures and Television Engineers (SMPTE) and the European Broadcasting Union (EBU). For experimental purposes, the use of existing formats has clear practical and economic advantages and has already enabled much progress to be made. However, it now seems likely that a new format will be necessary in order to realise the full potential of digital techniques.

Recent surveys[1] among users of videotape machines have established some of the features which will render digital machines commercially and operationally attractive. The elimination of noise, distortion and moiré, and the possibility of considerably increasing the acceptable number of copying generations made available by digital techniques must be evaluated.

This must be done in the context of such factors as editing, cassette operation and picture-in-shuttle, as well as of capital and operating costs. There remain several design problems; for example, editing and the inclusion of audio, as well as the agreement of

mechanical and track formats for scanning the magnetic surface, and data formats for organising the data patterns on the tracks. A digital VTR based on such an agreed standard is still several years distant from commerical availability.

Since 1976, the IBA has been active in promoting digital videotape recording as part of a continuing programme of work associated with the choice of digital video coding standards. This work has been well documented in previous volumes of the *IBA Technical Review* and in various published papers[2–9]. The work outlined in this article concerns the design and realisation of video recorders for coding standard proposals which require sampling of the separate component signals, that is, luminance (Y) and colour difference signals (B-Y) and (R-Y), rather than the composite signal.

Companion articles in this volume discuss the reasons for the preference for component coding and for the specific sampling frequencies proposed.

Two recorders are described: one based on a modified B format machine, the other on a modified C format transport. These machines were designed for different coding standard proposals. They continue to provide experience in operating different formats, as well as insight into their respective design approaches.

Factors Affecting Recorded Bit-rate

The choice of coding standards for digital video studios interacts with the design of appropriate DVTRs primarily through the bit-rate to be recorded. In earlier years, general opinion inclined to the view that the capabilities of the DVTR would be a major restriction on studio standards; but recent work has substantially proved otherwise. The bit-rates associated with the various proposals are shown in Table 1. The entries assume 8 bits per video sample.

Whereas the bit-rates are associated with the source coding processes, the data characteristics must be modified to match those of the recording channel. This can be achieved by means of an appropriate code transformation, herein later discussed. Further information must be added to the video information to enable correct identification of data on replay. Some impairment control information also may be recorded. These additions and the code transformation tend to increase the bit-rate which is to be recorded. This can be offset by suppressing the line blanking interval so that only the active line is recorded. A further small saving can be made by suppressing the vertical blanking interval at the expense of extra storage.

TABLE 1: SOURCE BIT-RATES CORRESPONDING TO THE VARIOUS SAMPLING FREQUENCY PROPOSALS CONSIDERED BY THE EBU

LUMINANCE SAMPLING FREQUENCY MHz	CHROMINANCE SAMPLING FREQUENCY MHz	SOURCE BIT-RATE Mbit/s
12	4	160
12	6	192
13	6.5	208
13.5	6.75	216
14.25	7.125	228

The nominal ratio of active line to total line period is 52 : 64 or 0.8125: but, with allowance for tolerances and for some horizontal picture shift, a ratio near 0.84 is chosen. Synchronisation information requires approximately 3% additional bit-rate, and impairment control can require a further 10–15%.

This latter factor is a relatively arbitrary choice, limited firstly by the need to minimise bit-rate, and secondly by the need of maintaining efficient control of impairment.

In total, the effective recorded bit-rate, R, is derived from the source bit-rate, S, as follows:

$$R = S \times \underset{\text{code conversion}}{1.25} \times \underset{\text{sync}}{1.03} \times \underset{\text{impairment control}}{1.125} \times \underset{\text{line blanking suppression}}{0.84}$$

$$R = S \times 1.22$$

This relationship is not definitive since the amount of line blanking suppression, synchronisation and impairment control information, as well as the code conversion ratio, may vary, perhaps considerably.

The calculated figure is therefore useful only as an approximation. If the IBA 8–10 code (to be described later) is not used, and a transformation involving no essential change in bit-rate is adopted, then it is possible to provide all the overheads within the savings achieved by line blanking suppression. This leads to a recorded bit-rate approximately equal to the source coding bit-rate. In this case, some additional impairment control allowance might be necessary to offset the loss of the error detection capabilities of the 8–10 code. This is because the codes

with low overheads have only small error detection ability.

For the machines to be described in this chapter, it may be assumed that, in general, the recorded bit-rate is 10–20% greater than the source bit-rate.

Bit-rate Reduction

To ease the recording of high bit-rates it is possible to use a bit-rate reduction scheme. A moderate reduction of, say, 20–30%, would be useful and is readily achieved by conventional methods. Bearing in mind that, in a studio, the highest possible quality must be maintained and that significant subsequent processing may be performed on a recorded signal, the removal of redundancy at such an early stage of production is undesirable.

Bit-rate reduction causes the valuable information content of a signal to be rendered more susceptible to impairment, roughly in proportion to the amount of redundancy removed. The possible need for the addition of error control information tends to neutralise the savings made by the redundancy removal process. The only advantage afforded by this arrangement is that, in removing natural redundancy from a signal and replacing it with a more systematic and controlled redundancy, the processes of impairment control may be rendered more efficient.

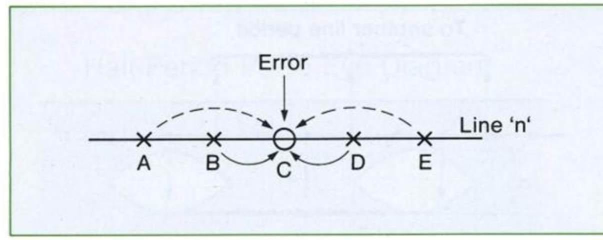
Impairment Control

The highly redundant nature of most video signals allows the concealment of an erroneous picture element by means of a substitute element derived from the picture elements which surround it (Fig. 1). These neighbouring elements are highly correlated over small regions of most pictures, and it is this correlation which manifests the picture redundancy.

A high degree of correlation can therefore exist between the true value of corrupted elements and predicted substitute values. This renders the substitution substantially invisible and subjectively acceptable.

Concealment is a convenient technique for dealing with errors without the aid of complex error correction codes. Such codes have limited capability of dealing with lengthy drop-out which occurs when head and tape momentarily lose contact.

There are many different techniques whereby such facilities could be provided by using a modest amount of additional recorded control information without the use of conventional error correcting codes. These codes are highly structured and can demand precise relationships with data formats. This could be



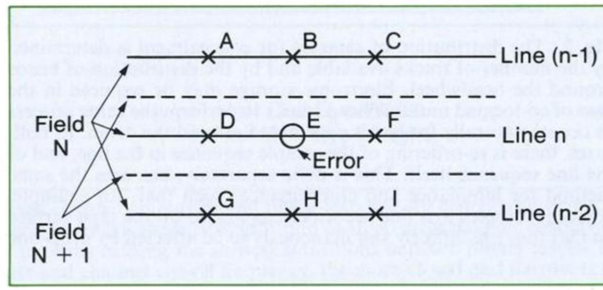


Fig. 1. The concealment of an erroneous sample may be achieved by using samples from the same line or from adjacent lines.

The first diagram shows the use of a single dimensional concealment using samples from the same line. By dispersing these samples over the tape surface, the effects of drop-out can be reduced, and thereby the concealment is substantially effective. Some simple algorithms for the replacement sample value Ĉ are shown; more complex ones requiring more hardware are possible.

The second diagram shows more complex, two-dimensional concealment using samples in the previous, same, and subsequent lines, together with some simple algorithms for the replacement sample. These samples are also distributed over the tape surface in an attempt to minimise the effects of drop-out.

Three-dimensional concealment is possible by sample substitution from previous fields or frames at the expense of the appropriate storage hardware. The concealment can, in all cases, and especially in the case shown in the first diagram, be adaptive so that erroneous samples are excluded from the calculation of the concealment. This allows choice of any preferred direction of concealment in sympathy with particular properties of the spatial spectrum near the concealment point.

inconvenient. The assistance of the recording channel code in the detection of erroneous data is of great value and is one of the factors influencing its design.

The control of picture impairment occurs in two stages: firstly, the detection of erroneous data; secondly, the replacement of this erroneous data by corrected data or by substitute concealment data. In either case, the detection of impairment is crucial since no control can be gained unless an indication of impairment is available. Also, false indications of impairment, or the breakdown of the capacity of the detection system in the presence of excessive error rates, as occurs during drop-out, can lead to extension of the influence of errors and must be avoided.

By means of 'scrambling' or 'shuffling' techniques, drop-out effects can be made to appear more like several small sequences or errors. Thereby, the

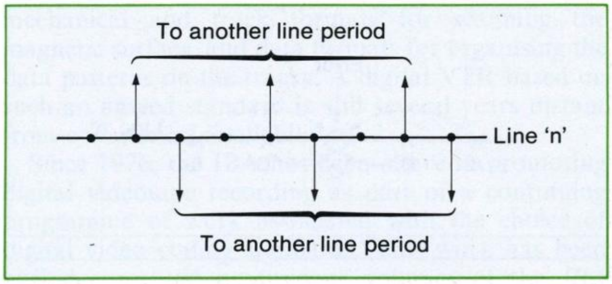


Fig. 2. The distribution of samples for concealment is determined by the number of tracks available and by the distribution of heads around the headwheel. Electronic storage may be required in the case of co-located multiple head stacks to perform the same process as occurs naturally for heads distributed around the drum. In both cases, there is re-ordering of the sample sequence in the line, and of the line sequence itself. This is done separately, but uses the same method for luminance and chrominance; such that, for example, adjacent samples in a line are widely separated on the tape surface so that they are unlikely simultaneously to be affected by drop-out.

natural sequence of bits, samples or lines produced at the source is re-ordered by means of storage.

This is a powerful means of enabling simple concealment techniques to control quite severe drop-out effects. The concealment methods tend to be particular to mechanical and data format designs; the detailed arrangements, corresponding to the B format and C format machines developed at the IBA, are described later.

Code Conversion

The matching of source data characteristics to those of the recording channel requires the use of some code transformation. The presence of rotary transformers as part of the scanning mechanism, and the lack of d.c. and low frequency response in the magnetic recording medium, require the suppression of spectral energy in these regions. The source data is usually coded as non-return-to-zero binary (NRZ), the power spectrum of which is of the form:

$$\left[\frac{\sin x}{x}\right]^2$$

and, clearly, this has a d.c. component and substantial power in the low frequency region. Further, at the frequency corresponding to the bit-rate, there is a spectral null, and non-linear processing is required to generate a clock component. The amplitude and phase modulation of this clock are caused by the data pattern variations. They can be reduced by careful channel response shaping and by ensuring that the code transformation provides data transitions at frequent intervals.

There are numerous code transformations which will perform the required modification. One which has been developed at the IBA, and used under licence by other experimenters, involves the representation of 8-bit video samples by 10-bit numbers. Of the $2^{10} = 1{,}024$ possible 10-bit numbers, 252 have the property that they contain equal numbers of '1' and '0' and therefore have no inherent d.c. component. Further, sequences of these numbers will have suppressed low frequency spectra, as compared with those of NRZ. The fact that there are five '1's and five '0's in each number guarantees transitions no further apart than 10 bits, this worst case being a very rare occurrence.

The use of only 252 permissible states with 256 8-bit combinations requires the suppression of four of these combinations. The precisely limited dynamic range of the analogue-to-digital converters (ADC) requires the maintenance of adequate headroom for the analogue video signals. The loss of two quantisation steps at the extremes of the ADC range is therefore acceptable and can be considered as part of the system headroom.

One apparent disadvantage of the 8–10 code conversion is that a 25% increase in bit-rate is produced. However, because there are $1{,}024 - 252 = 772$ unused 10-bit numbers, the occurrence of these groups in recovered data indicate

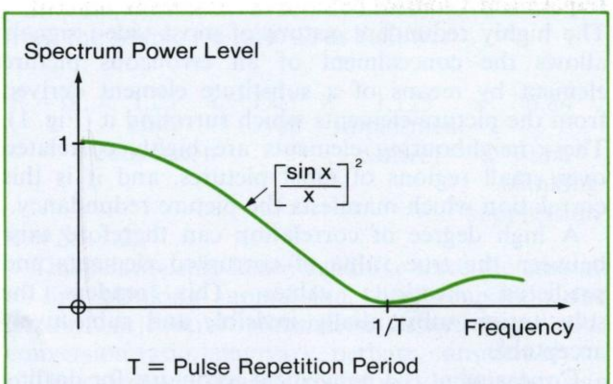


Fig. 3. The power spectrum of the binary, non-return-to-zero (NRZ) data format is of the form

$$\left[\frac{\sin x}{x}\right]^2.$$

It has considerable density near d.c. and a null at a frequency of period equal to the NRZ bit period. For a signal amplitude of unity, the spectrum also contains at d.c. a line of one-half amplitude.

For a random data stream, the spectrum is substantially noise-like spectral lines being generated only by repeating sequences of bits.

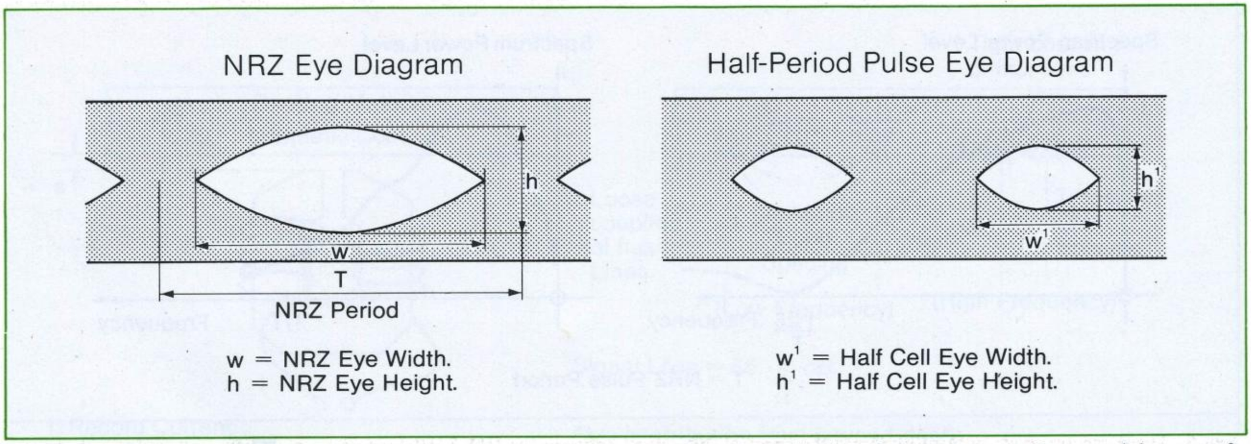


Fig. 4. The shape of the impulse response can be calculated for any given channel amplitude and phase response. In general, however, for any random sequence of bits, a more practical indicator is the 'eye' diagram which is the superimposition of all possible data sequences over a single bit period. The 'eye' opening is defined by the worst case bit patterns causing the slowest transitions between binary levels. The more severe the rate of cut-off, and the higher the ratio between bit-rate and channel cut-off frequency, the more closed and narrow is the eye.

A typical eye opening is shown here for an NRZ data format compared to that for a coded signal which uses additional transitions halfway between those of the NRZ signal. The additional central transition causes the NRZ eye to be split into two smaller 'eyes' which, for the same channel shaping, are half the height ($h^1 = h/2$) and less than half the width ($W^1 \leqslant W/2$) of the NRZ case. This can lead to greater instrumental complexity in regenerators, which are required to recover the data sequences from these degraded 'eyes' compared to the NRZ case. This is because of the greater amplitude resolution required and the smaller timing jitter tolerance available to meet given error rate criteria.

the presence of bit errors. This property is useful; and the 25% overhead of the code can be viewed partly as an impairment control allowance. The scope of this error detection capability is limited since those errors which cause the transformation of one valid 10-bit group into another are undetectable. The 8–10 code can detect all odd numbers of bit errors in each word, but only some of the even numbers; for example, only 50% of 2-bit errors.

The 8–10 code is readily realised by means of Programmable Read Only Memories (PROM) where the input address is the 8-bit video sample and the corresponding contents are the appropriate 10-bit number. Decoding is performed by a PROM having 10-bit input addresses with contents corresponding to the appropriate 8-bit video sample values. The serial bit cell width is 25% narrower than for the NRZ source data. Compared with some other codes, which place coded transitions in the centre of the NRZ cells, the 8–10 code requires less bandwidth; and it provides eye diagrams which are more readily resolved (Fig. 4).

Some codes which employ sub-division of the bit cell can be arranged to suppress those combinations of transitions leading to high frequency spectral energy, while also suppressing d.c. and low frequency energy (Fig. 5).

This ability is at the expense of other features such as: (*a*) ease of realisation, that is, the encoding and decoding circuits can be complex; (*b*) degraded eye diagrams, due to the cell sub-division; and (*c*) error detection capabilities.

It is believed that the 8–10 code provides a good compromise between the various conflicting requirements of code design.

The question of code conversion is still a relatively open one and, when the standardising of a recording code is considered, all the above mentioned factors must be taken into account.

Features of Digital Operation

It is generally the case that, when digital techniques are employed, signal processing of the digital waveform, especially in its binary form, demands more bandwidth than does its analogue counterpart.

Techniques are available to reduce this bandwidth by an almost arbitrary amount; either by redundancy reduction at the source, or by transformations using multilevel codings such as those currently common in digital transmission. In these latter cases, the bandwidth reduction is obtained by reducing channel symbol repetition rates to values less than the source bit-rate. It is available by the use of a substantially linear channel affording a trade-off with signal-to-noise ratio snr. In magnetic recording channels this linearity is not available.

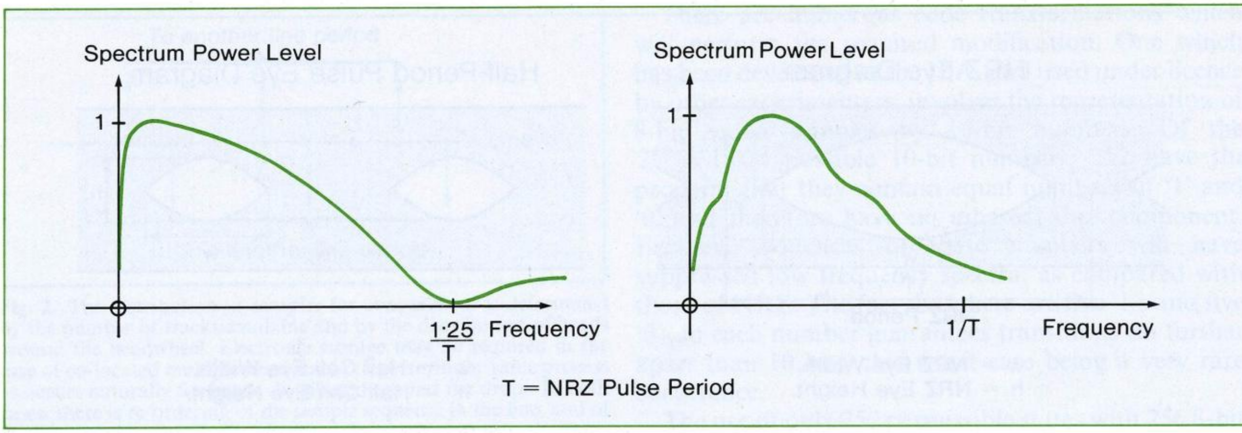


Fig. 5. These spectral distributions illustrate the spectral differences between the IBA 8–10 code and a code using bit period sub-division (Miller² Code). In each of these the lack of low-frequency energy is notable, as is the extension of the Miller² energy above the frequencies corresponding to the NRZ bit period. The IBA code has a spectral null at 1.25 times this frequency. However, in practice, both signals would be subject to band-limiting from below this frequency in order to control noise and thus optimise snr.

The digital binary waveform is the most rugged format and can function adequately for broadcast video purposes at baseband snr values near 20 dB. Multilevel coding is not readily applicable to magnetic recording due to the bistable nature of the medium, and to its non-linearity.

Ternary coding is possible if the unmagnetised state of the medium is permitted as an extra state; but, in practice, this state is highly vulnerable to corruption by stray magnetic fields and is somewhat difficult to arrange.

One particular advantage of digital operation over analogue is the convenience of time division multiplexing the information into several relatively low bit-rate data streams, each of which can be associated with a separate channel and track. This arrangement is less readily achieved in the analogue case although, by frequency division multiplex of the analogue signal, a similar result is possible. Distribution of information among several tracks clearly reduces the bit-rate per track. Also, it obviates the need for a very high relative speed between head and tape when the wideband digital signal is recorded on a single track.

In digital video recording, adequate bandwidth and snr must be made available to support the bit-rates demanded by the picture source. The parameters of bandwidth and snr are closely related to wavelength in the medium. The bandwidth affects signal fidelity as well as noise level, while the snr is connected with impairment due to bit errors. In recording, an additional factor, also connected with wavelength, is important. It is the condition arising from the well-known Wallace Effect[10] which leads to drop-out and, hence, to error bursts.

Drop-out occurs when head and tape momentarily lose contact, such that a high reluctance air-gap separates them, thus reducing the coupled flux. This reduction is heavily wavelength dependent; the empirical result due to Wallace suggests that the flux loss caused is 55 dB per wavelength of separation. Thus, in practice, the shorter wavelengths suffer from drop-out phenomena more than do longer wavelengths. However, this basic effect is active even when head-to-tape contact is close. Clearly, flux contributions from within the magnetically active bulk of the medium are reduced compared to contributions from the surface.

This is especially the case at short wavelengths since the flux is related to the number of active particles; and so, the flux per bit will be directly related to volume. Therefore, if recording at short wavelengths is attempted, the volume of active material associated with each bit is reduced in proportion to the square of the wavelength.

Thus, replay signal level for a given track width is proportional to the squared wavelength; and so, other things being equal, the trade-off between snr and bandwidth can be effected by reducing recorded wavelengths in order to increase bit-rates.

The number of magnetic particles per bit is also proportional to track width; so that, reducing snr by reducing track width allows exploitation of a further dimension of trade-off with bandwidth.

For any given tape consumption, the number of independent tracks can be changed, provided that the

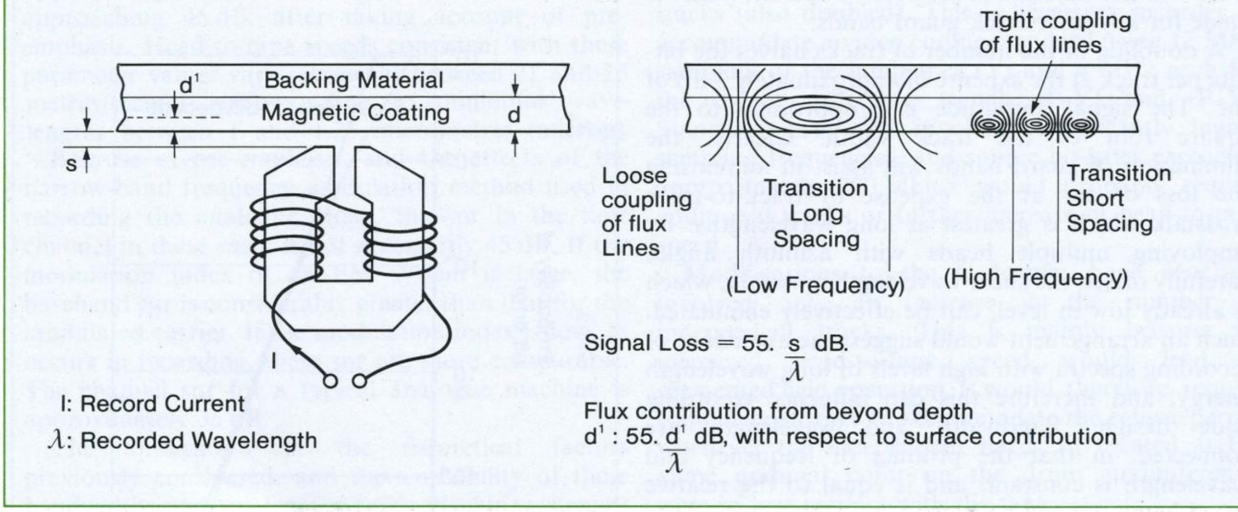


Fig. 6. The optimal recovery of a signal from a recorded tape requires intimate magnetic contact between head and tape in order to maximise snr and to minimise impairment. Any separation between head and tape causes a wavelength dependent loss, as here indicated. Even with smooth tape surfaces, the effects of dust and trapped air between head and tape can so reduce the coupled flux that drop-out effects can occur. If head and tape contact were perfect, this effect would still be in evidence; because, at short wavelengths, the flux contributions from deep in the tape are reduced according to the same rule as for head-tape separation. In addition, short wavelengths involve close packing of effective magnetic poles which maintain tight flux coupling so that recovered signal levels are reduced. Equalisation of the wavelength response has a noise penalty at high frequencies and, overall, the snr is heavily dependent on wavelength.

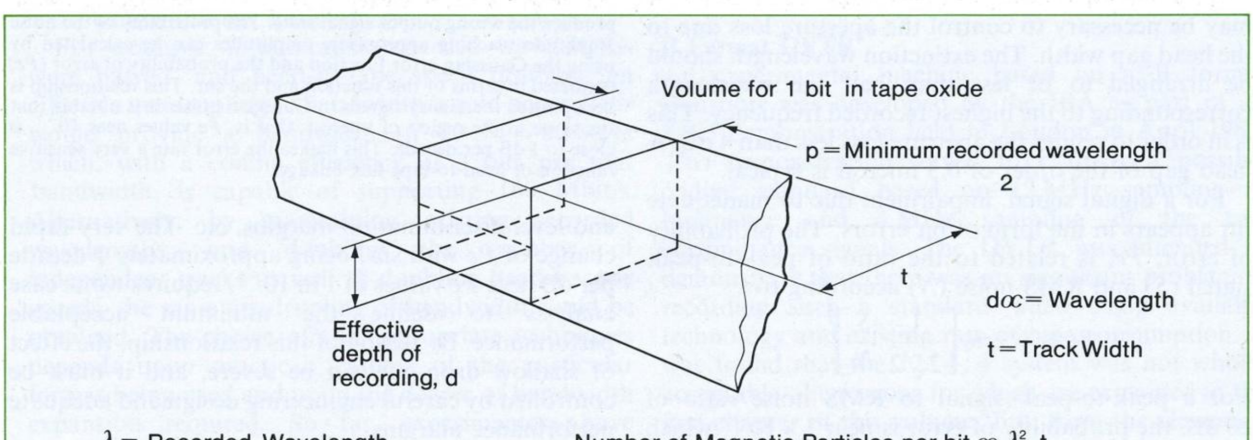


Fig. 7. The effective signal level recovered from the tape depends on the number of magnetic particles contributing to the magnetic circuit flux. The number of particles contributing flux is a function of wavelength, because the depth of recording is wavelength dependent due to the spacing loss effect. Along the track, the distance occupied by each bit is half the minimum recorded wavelength. Therefore, the equivalent volume of each bit is $t \times p \times d$, where t is track width, p is bit-length along track, and d is mean recording depth. Both p and d are related to minimum recorded wavelength; and so, snr is a square function of wavelength. This can be exploited to enable the trade-off of reduced snr for shorter wavelengths.

track width is adjusted in proportion and allowance is made for any inter-track guard bands.

A doubling of the number of tracks halves the bit-rate per track at the expense of approximately 3 dB of snr. The signal amplitude is proportional to the square root of the track width. Clearly, the elimination of guard bands will assist in minimising the loss of snr at the expense of track-to-track crosstalk which is greatest at long wavelengths. By employing multiple heads with azimuth angles carefully offset, the short wavelength crosstalk, which is already low in level, can be effectively eliminated. Such an arrangement would suggest the avoidance of recording spectra with high levels of long wavelength energy, and therefore this can influence recording code design. Bandwidth and wavelength are connected, in that the product of frequency and wavelength is constant, and is equal to the relative speed between head and tape.

For a fixed value of head-to-tape speed, the trade-off between wavelength and bandwidth, already described, can be exploited. By increasing head-to-tape speeds, wavelengths are proportionally increased, thus giving snr improvement which is related to the square of wavelength increase. Thus, an increased head-to-tape speed can also be used to increase bandwidth and, hence, bit-rate. In general, the bandwidth is given by head-to-tape speed divided by minimum wavelength.

If the recorded head-to-tape speeds are modified, it may be necessary to control the aperture loss due to the head gap width. The extinction wavelength should be arranged to be less than half the wavelength corresponding to the highest recorded frequency. This is in order to render the aperture loss less than 4 dB. A head gap of the order of 0.3 micron is typical.

For a digital signal, impairment due to inadequate snr appears in the form of bit errors. The probability of error, *Pe*, is related to the ratio of peak-to-peak signal (S) and RMS noise (N) according to:

$$Pe = \tfrac{1}{2}.\,\mathrm{erfc}\left[\frac{1}{2\sqrt{2}}.\frac{S}{N}\right]^{\frac{1}{2}}$$

For a peak-to-peak signal to RMS noise ratio of 20 dB, the probability of error is near 1×10^{-7} which is virtually undetectable in broadcast standard video signals, provided the errors remain statistically random.

Ideally, therefore, the DVTR snr should be in excess of 20 dB to allow some margin for practical realisation factors such as tracking accuracy, equalisation response variations, regenerator timing and level discrimination margins, etc. The very rapid change of *Pe* with snr, being approximately 1 decade per dB near *Pe* values of 1 in 10^{-6}, requires worst case analysis to define the minimum acceptable performance. By reason of this relationship, the effect of shallow drop-out can be severe, and it must be controlled by careful engineering design and adequate performance margins.

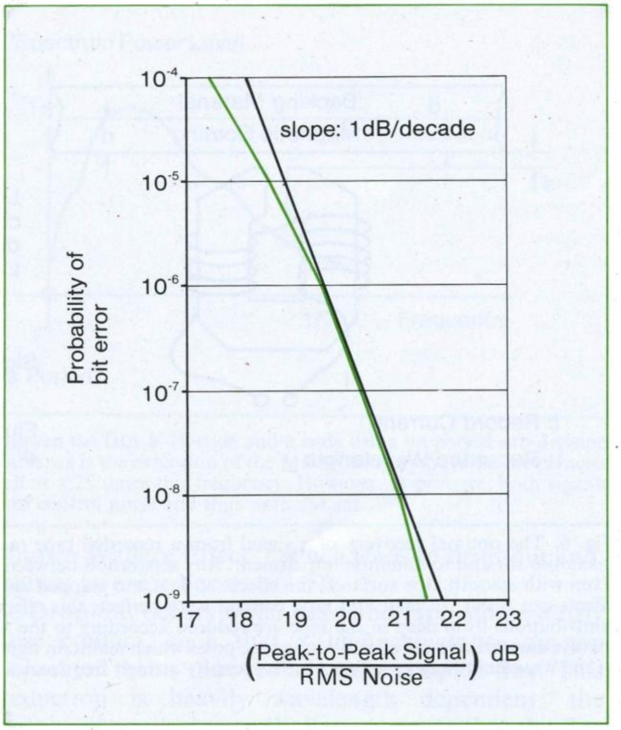


Fig. 8. The probability of error is related to the peak-to-peak signal to RMS noise ratio in the channel. The noise is usually assumed to be Gaussian and stationary. An error is assumed to be caused when the peak signal amplitude is exceeded by a noise peak during regeneration of a bit. This causes the resultant signal amplitude to fall the wrong side of the regenerator amplitude threshold and so produce the wrong output signal state. The probability of the noise amplitude reaching appropriate amplitudes can be calculated by using the Gaussian error function and the probability of error (*Pe*) predicted in terms of this function and the snr. This relationship is here plotted for binary signals in Gaussian noise. It is notable that the slope in the region of interest, that is, *Pe* values near 10^{-7}, is close to 1 dB per decade. This makes the error rate a very sensitive function of head-to-tape flux linkage.

Features of Experimental DVTR

Digital operation has been achieved by experimental modification of existing analogue transports. Although somewhat dependent upon head and tape characteristics, the short wavelength performance of current analogue VTRs is adequate for bandwidths

approaching 20 MHz, and baseband snr approaching 45 dB, after taking account of pre-emphasis. Head-to-tape speeds consistent with these parameter values vary somewhat between 21 and 25 metres/second corresponding to minimum wavelengths between 1 and 1.25 micrometres (micron).

Because of pre-emphasis, and the effects of the narrow-band frequency modulation method used in recording the analogue video, the snr in the tape channel in these cases is not necessarily 45 dB. If the modulation index of an FM system is large, the baseband snr is considerably greater than that for the modulated carrier. If the modulation index is low, as occurs in recording, these snr are more comparable. The channel snr for a typical analogue machine is approximately 35 dB.

The application of the theoretical factors previously considered, and the availability of these bandwidths and snr from existing machines, suggest various methods of applying high-speed digital recording. It is possible to use for this purpose the following three techniques, either singly or in combination:

(*a*) reduce minimum recorded wavelengths to enable a bandwidth increase at the expense of snr;

(*b*) increase the number of independent tracks so that each track retains existing bandwidths but the reduced track width leads to reduced snr;

(*c*) increase head-to-tape speed to increase recorded wavelengths and/or bandwidths, still allowing a bandwidth/snr trade-off.

For example, if minimum recorded wavelengths were halved, and head-to-tape speed doubled, an effective quadrupling of bandwidth would occur. This would provide a bandwidth of $20 \times 4 = 80$ MHz which, with a coding efficiency of 2 bits per unit bandwidth, is capable of supporting 160 Mbit/s. Alternatively, by maintaining existing recorded wavelengths, and doubling the number of independent tracks as well as doubling head-to-tape speeds, the same quadrupling of bandwidth could be obtained. The choice of the appropriate techniques depends upon practical features of the particular format being used and upon the degree of bandwidth expansion required. So far, experimenters have tended to maintain, or even increase, minimum recorded wavelengths in the region of 1–2 micron. This leads to the need to increase the number of independent tracks or to increase the head-to-tape speed, or both.

Modifications to the B format have involved the use of increased head-to-tape speed (approximately doubled), and an increased number of independent tracks (also doubled). This is necessary in order to accommodate a video coding standard using 12 MHz sampling of the luminance (Y) and 4 MHz each for the colour difference signals (B-Y) and (R-Y). Standards proposals involving significantly higher sampling frequencies, and source bit-rates exceeding approximately 200 Mbit/s would probably require additional tracks or further increase of head-to-tape speed.

Modifications to the C format have generally involved only an increase in the number of independent tracks. This is mainly because an increased head-to-tape speed would lead to segmented field operation. It would, therefore, require some data storage to accommodate the eclipse period whenever the independent heads (if located at the same nominal point on the drum circumference) were to lose contact with the tape.

Segmented field operation is perfectly feasible with C format transport; and the choice of method of digital conversion is a matter of practical convenience as well as of theoretical interest. It is also possible to locate the independent heads at points spaced around the circumference of the drum; this has some advantages over the co-location of heads, but there are also disadvantages.

The details of the B and C format DVTRs developed at the IBA are discussed separately in the following two sections.

B Format DVTR

An experimental machine based on a B format transport was developed by the IBA as part of an EBU demonstration held in London in April 1980. This demonstration served to evaluate a possible coding standard based on 12 MHz sampling of luminance and 4 MHz sampling of the two chrominance signals. The DVTR was intended to demonstrate that there was no significant problem in recording such a standard when using available technology and existing rate of tape consumption. It was found that the 12 : 4 : 4 system was not wholly acceptable, the reasons for which are explained in the first chapter of this volume; but, from the viewpoint of the DVTR, some processing techniques were shown as being useful and some features of the B format shown as having practical advantages.

The 12 : 4 : 4 system generates 160 Mbit/s at source. With the addition of synchronisation information, and the use of the IBA 8–10 channel code, a recording bit-rate of 175 Mbit/s was chosen.

Line blanking was suppressed, but no recorded error control information was provided. The resultant bit-rate is derived as follows:

175 = 160	× 0.85	× 1.25	×	1.03
Source rate	line blanking suppression	8–10 code		housekeeping

At a coding efficiency of 2 bits per unit bandwidth, this rate would require a channel bandwidth of 87.5 MHz for a single track format. This, in turn, would require a head-to-tape speed of 87.5 metres/second for a minimum wavelength of 1 micron. As this head-to-tape speed would require impractical headwheel revolution speeds, two independent tracks were chosen, the second track being provided by an additional head pair mounted opposite each other on a diameter at right angles to the original, similarly mounted, head pair.

The information can be multiplexed conveniently between the two tracks in order to assist in impairment control, which will be discussed later. Thus, each track now accommodates 87.5 Mbit/s and requires a bandwidth of approximately 45 MHz which can be provided with acceptable wavelengths by a head-to-tape speed of 50 metres/second. This head-to-tape speed is practicable and is almost exactly double the value used in the analogue B format case. A headwheel angular velocity of 312.5 Hz was selected because this frequency is exactly one-fiftieth of line frequency for the 625-line 50 Hz scanning standard. Thus, exactly 50 lines can be recorded per headwheel revolution. This has advantages in signal processing and head switch synchronisation.

The data is distributed between the tracks such that Y, U and V samples each alternate between them. It is clear that, when these samples are restored to their correct places in each recorded picture line, every alternate Y, U and V sample is recovered from a different track. Further, the use of head pairs located at points equally spaced around the headwheel circumference means that, while one head pair is beginning a scan and its partner ending a scan, the other head pair is halfway through a scan, i.e. one head in contact, the other not in contact with the tape. This arrangement means that alternate Y, U and V samples are placed on the tape at one half track length intervals.

This ensures that any tape blemish or drop-out affecting one head pair is unlikely to be affecting the other. By this means, a simple single dimensional interpolation from adjacent samples in the line can be used to provide an alternative sample to an erroneous one.

Audio information is provided for in this machine by arranging for exploitation of the overlap period where both heads of a pair are simultaneously in

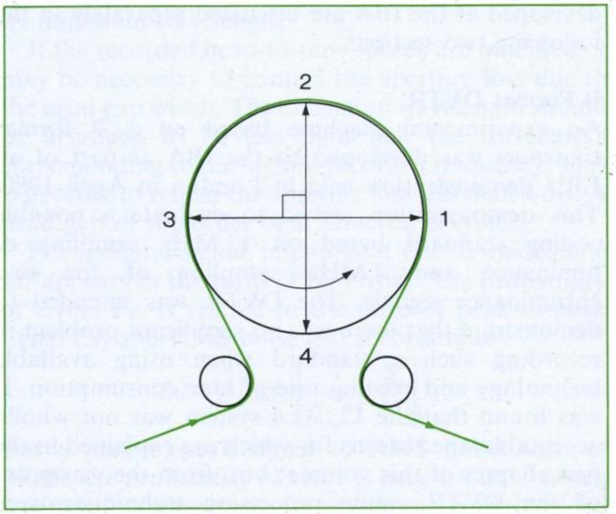


Fig. 9. This illustrates the distributed arrangement of heads for the modified B format experimental system. This has the advantage that, for example, if head 1 begins its scan recording alternate samples along any given line, then the remaining samples, simultaneously routed to head 2, are recorded half-way along a scan, and thereby automatically become physically separated on tape. If the heads were co-located, this physical separation of data would need be arranged by means of electronic storage.

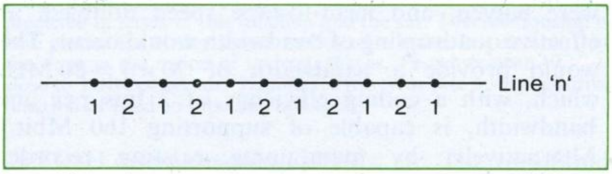


Fig. 10. This illustrates the distribution of data samples required to exploit the head arrangement shown in Fig. 9.

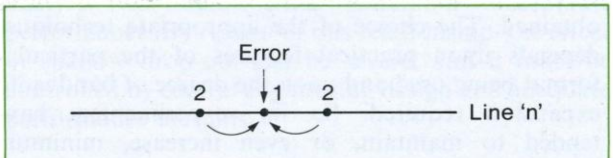


Fig. 11. This illustrates the simple concealment algorithm used for protection against errors. The arrangement of Fig. 10 ensures the physical separation of adjacent samples; and the head arrangement of Fig. 9 ensures that these samples are recorded at half-scan intervals along adjacent tracks. The probability of any tape blemish or drop-out affecting both sets of samples is therefore remote. Thus, high quality concealment is achieved by simple means.

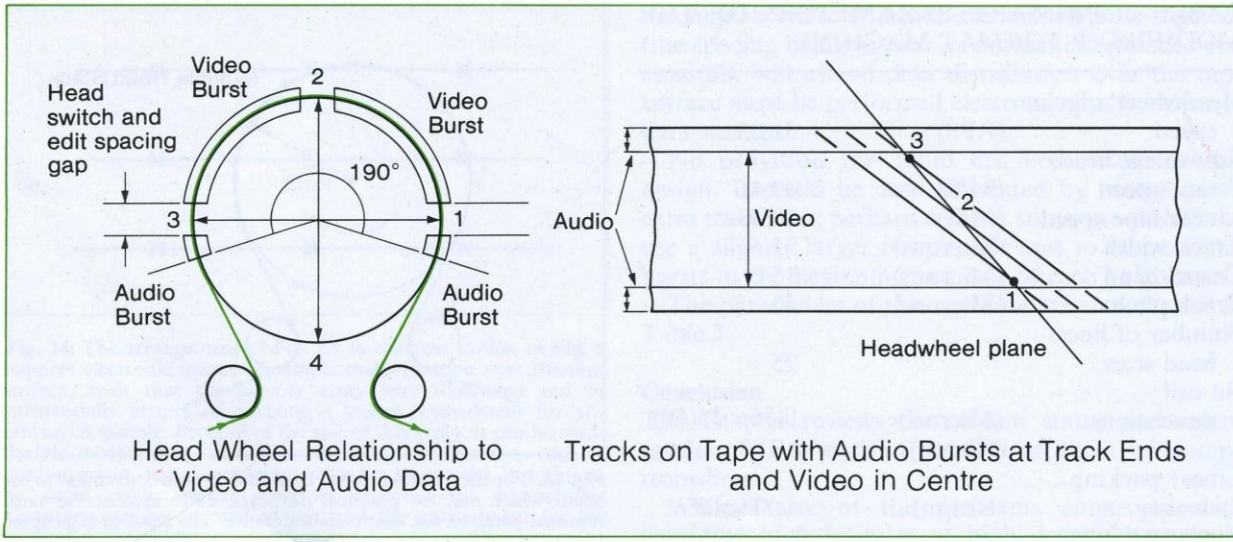


Fig. 12. The arrangements of heads shown in Fig. 9 conveniently combines with the wrap angle of 190° to enable the recording of audio and video as here illustrated. The overlap period of 10° for each head pair allows 5° before and after each head switch instant for audio data. Also, it exploits the period when both heads of a pair are simultaneously in contact with the tape. Regular line related head switches are possible because of the synchronisation of the headwheel rotation speed to line frequency (312.5 Hz = line frequency ÷ 50). A dummy head switch, half-way down each video scan, allows synchronous recording and replay of the two channels of data. This is convenient and serves to simplify the hardware. The corresponding track arrangement, and the location of the audio and video data on the tape, are here illustrated.

contact with the tape. This requires that a gap be created between the audio and video sectors of the track in order to gain separate access for editing purposes. In addition, to facilitate head-switching when headwheel tolerances and tape interchange may be involved, a short gap is created in the video data coincident with the headswitch instants. In order to create these gaps, the effective line period on tape is slightly reduced. The video is confined within the central 180° of the 190° wrap angle, and the audio occupies the remaining 5° at either end of the track.

The machine provided acceptable picture quality despite the relatively simple error control mechanism employed. The advantages for impairment control, of using heads distributed around the headwheel, are clear, as is the use of the overlap period for audio. The recorded wavelength employed is 1.14 micron; and this serves to limit all drop-out phenomena to an acceptable level.

Synchronisation information is recorded as a block preceding every half-line of picture data. The more frequent the synchronisation information, the better the control over system timing, but at the expense of additional bit-rate. Although this is a relatively small overhead, it is an important one, because the loss of a synchronisation indication, say, due to errors, could lead to incorrect decoding. Thus, a complete block of correct picture data could be lost, and the subjective picture quality be severely impaired. The parameters of this machine are summarised in Table 2.

C Format DVTR

An experimental machine based on a C format transport was provided by the IBA as part of an EBU demonstration held at Crawley Court, Winchester in January 1981. This demonstration involved evaluation of several proposed standards for sampling. The DVTR was intended to demonstrate the feasibility of recording bit-rates associated with the highest proposed sampling frequencies; it was also designed as to be easily adaptable for operation either on 625-line 50 Hz or 525-line 60 Hz scanning standards.

The highest sampling frequency proposals involved 912 luminance samples per line and 456 samples each for the colour difference signals. The bit-rate generated by this arrangement is 228 Mbit/s. With the addition of synchronising information and the use of the IBA 8–10 code, a recording bit-rate of 277.5 Mbit/s is obtained. This allows some capacity for an error control system, in addition to that provided by the 8–10 code. It requires the recording of control signals along with data so that, on replay, these control signals can be recovered and used for

TABLE 2: PRIMARY PARAMETERS OF THE MODIFIED B FORMAT MACHINE

Headwheel angular speed	(RPS)	312.5
Equivalent head/tape speed	(M/S)	49.41
Linear tape speed	(M/S)	0.243
Track width	(Microns)	35
Guard band	(Microns)	15
Track pitch	(Microns)	50
Number of lines/head scan		25
Bit cell dimensions	(Micron × Micron)	$0.56 \times 35(50)$
Linear packing density	(Bits/m)	1.77×10^6
Linear packing density	(Bits/inch)	4.5×10^4
Area packing density	(Bits/M²)	3.5×10^{10}
Area packing density	(Bits/inch²)	22.85×10^6

detection of errors.

At a coding efficiency of 2 bits per unit bandwidth, this rate would require a channel bandwidth of 139 MHz for a single track format, and an associated 139 metres/second head-to-tape speed, for a minimum wavelength of 1 micron.

The C format transport operates in its analogue mode such that the heads are in contact with the tape for 300 lines per field; the wrap angle is approximately 346°.

The increasing of head-to-tape speed would lead to field segmentation and, with a single record or replay head, would therefore require storage to accommodate the eclipse period. The solution could be to distribute heads around the drum; but, with the large wrap angle, this could not be done conveniently without some loss of efficiency of tape contact. The solution adopted in the case of C format is therefore to retain non-segmented operation and the head-to-tape speed of normal analogue C format, and to employ a multiple head stack. This allows the simultaneous recording, or replay, of several parallel tracks. Thus, the head-to-tape speed is approximately 21.2 metres per second for the 625-line 50 Hz scanning standard, and 25.4 metres per second for the 525-line 60 Hz scanning standard.

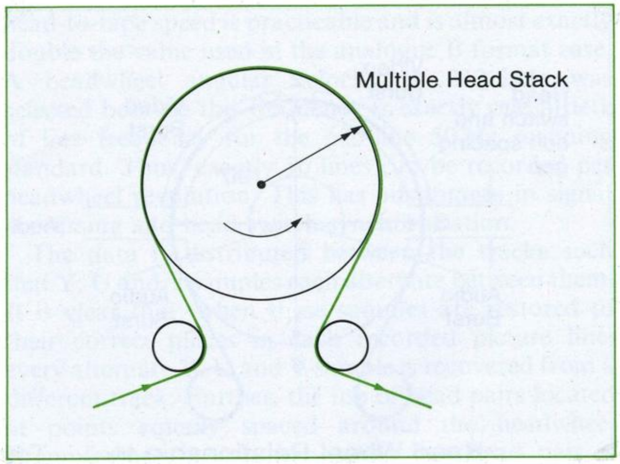


Fig. 13. For the C format it has been found more convenient to use heads which are, for practical purposes, co-located at the same nominal point on the drum circumference. The head-to-tape speed is unchanged from the analogue case.

In the case of 525-line, this 20% increase in head-to-tape speed results in 20% increase of wavelengths, to reduced spacing loss, and reduced drop-out effects. The bandwidths available at these head-to-tape speeds are in the region of 20–25 MHz; therefore, they can support 40–50 Mbit/s at 2 bits per unit bandwidth coding efficiency. Thus, a machine designed for 277.5 Mbit/s must have six or seven separate tracks. In practice, six heads were employed, each one recording and replaying at 46.25 Mbit/s. A greater number of independent heads would allow reduction in bit-rate per channel or increase in machine capacity. The accommodation of multiple tracks requires that the individual track widths be reduced, otherwise increased tape consumption would be required. The track widths employed are in the region of 30 microns.

In order to combat the effects of impairment, the picture data are dispersed in time and, therefore, in space on tape. As for the B format machine already described, concealment techniques are employed. In this case, not only are picture samples within each line re-distributed, complete lines are re-sequenced so that any consecutive picture lines are several lines apart on tape. This, together with a two-dimensional concealment arrangement, enables the concealment system to overcome lengthy drop-outs by rendering them similar to single sample errors, and thereby improving the facilities provided for the B format machine. In the B format machine, this distribution was possible by virtue of the distribution of the heads around the headwheel circumference. In the C format

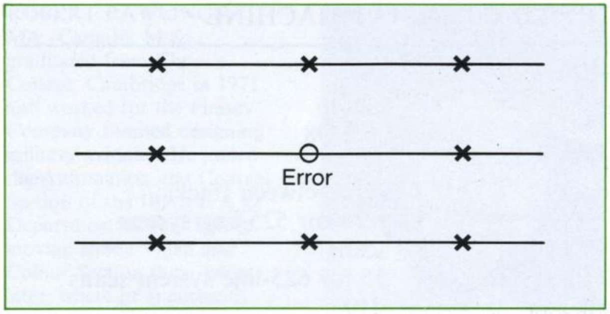


Fig. 14. The arrangement of Fig. 13, in contrast to that of Fig. 9, requires electronic means of sample re-distribution over the tape surface, such that the sample array here illustrated can be substantially certain of enabling a sound concealment for any erroneous sample. Because of the size of this array, it can be made flexible in operation, such that erroneous contributions (due to random errors) to a concealment value can be ignored. In addition, by some crude spatial spectrum analysis, the concealment algorithm can be adapted to optimise the concealment for the particular spatial frequencies present.

machine, because all heads are located close together (the spacing being as near as mutual interference and crosstalk will allow) data distribution over the tape surface must be performed electronically by means of data storage.

No provision for audio has been made in this design. It could be accommodated by means of an extra track; but, perhaps a better solution would be to use a slightly larger wrap angle and to insert audio bursts at the beginning and end of each video scan.

The parameters of this machine are summarised in Table 3.

Conclusion

This chapter reviews the current status and recent results of IBA work in the field of digital videotape recording.

While many of the problems connected with recording high bit-rates at high density have been

Plate. This off-screen photograph of a digital videotape recorder playback was taken during the EBU demonstrations in January 1981. Note the excellent picture quality available.

TABLE 3: PRIMARY PARAMETERS OF THE MODIFIED C FORMAT MACHINE

Headwheel angular speed	(RPS)	50
Equivalent head/tape speed	(M/S)	21.146
Linear tape speed	(M/S)	0.2398
Track width	(Microns)	29
Guard band	(Microns)	zero between tracks zero for 525-line system scans 35 for 625-line system scans
Track pitch	(Microns)	210
Number of lines/head scan		288
Bit cell dimensions	(Micron × Micron)	0.45×29
Linear packing density	(Bits/m)	2.2×10^6
Linear packing density	(Bits/inch)	5.59×10^4
Area packing density	(Bits/M^2)	6.28×10^{10}
Area packing density	(Bits/inch^2)	41×10^6

solved, there remain several important problems such as picture-in-shuttle and editing. These operationally important features are closely connected with the design of a suitable mechanical format for digital recording as well as data formats.

Future work in the solving of these remaining problems is likely to lead to the introduction of a fully operational digital videotape recorder.

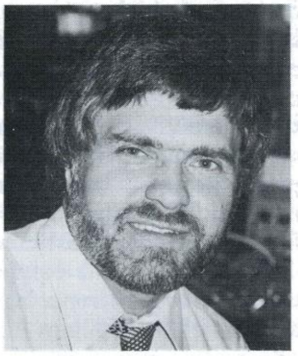
Acknowledgement
The work described in this chapter was performed at the IBA by numerous contributors; the efforts of Richard Morcom, Alastair Angwin and Christopher Birch are particularly appreciated for successfully completing hardware design and realisation in very limited time scales. In addition, the assistance of industry for information and the supply of tape transport hardware for experimental purposes is gratefully acknowledged.

References

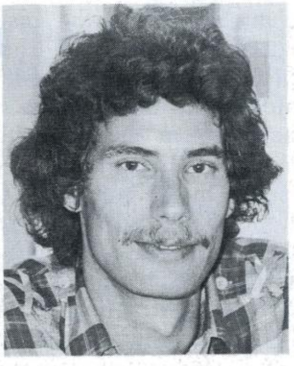
1. P. N. Kelly, 'User Requirements of Digital Video Tape Recorders', *IEE Conference Publication No. 191, IBC '80*, 227.
2. *IBA Technical Review 12.*
3. J. H. Wilkinson and K. Lucas, 'High Quality Transmission of Digital PAL Signals at 34 Mbit/s', *IEE Conference Publication No. 166, IBC '78*, 42.
4. J. L. E. Baldwin, 'Digital Television Recording with Low Tape Consumption', ibid, 133.
5. J. L. E. Baldwin, 'Digital Coding Standards for Television Studios', *IEE Conference Publication No. 191, IBC '80*, 85.
6. E. J. Wilson, 'A Digital Codec for Satellite TV', ibid, 304.
7. G. M. Drury and J. L. E. Baldwin, 'Digital Video Tape Recorders—Formats and Standards', ibid. 236.
8. J. L. E. Baldwin, 'Codes for Digital Video Tape Recording at 10 Mbit/square inch', *IERE Conference Proceedings No. 43*, Video and Data Recording Conference (University of Southampton, 24–27 July 1979), 147.
9. G. M. Drury, 'A Tape Interface System for High Density Digital Recording', ibid, 241.
10. R. L. Wallace, 'Reproduction of Magnetically Recorded Signals', *BSTJ Vol. 30* (October 1951), 1146–73.

Chromakey in Future Studio Systems

by R. Rawlings and N. Seth-Smith



ROBERT RAWLINGS, MA (Cantab), M.Sc., graduated from Christ's College, Cambridge in 1971 and worked for the Plessey Company Limited designing military avionics. He joined the Automation and Control Section of the IBA's E & D Department in 1974 before moving to the Video and Colour Section three years later, where he is currently Principal Engineer responsible for investigating international studio standards. Married with a son and daughter, he lives in Hampshire and enjoys sailing and playing football and tennis.



NIGEL SETH-SMITH, B.Sc., joined the IBA as a graduate trainee after obtaining a degree in electronic engineering at Southampton University. At the end of the two-year training course he worked for one year in the Transmitter Section of the Station Design and Construction Department. He then left the IBA to work as part of a small team designing broadcast quality television cameras with Link Electronics Limited. In 1978 he returned to the IBA to join the Video and Colour Section where he has been involved with work in connection with the choice of a studio standard for digital coding. He is married, lives in Hampshire, enjoys skiing and is a keen tennis player.

Synopsis
Chromakey is typical of the complex signal processing necessary in modern television studios, and its use as a production technique is of growing importance. Much discussion about digital studio sampling standards has centred on the colour channel bandwidth requirements. The effect that these will have on digital chromakey performance is here discussed. Two experimental investigations are described, namely the design of a digital chromakey equipment and a study of the relationship between sampling standards and chromakey performance.

Introduction

The technical performance of chromakey has been developed to such an extent that its use has risen above the level of gimmickry to become an everyday occurrence in television broadcasting. In essence, the process involves the use of a highly saturated area of colour in one television picture to control an electronic switch, or 'key', to a different source. The result is thus an electronic composite in which the areas of picture containing the saturated colour have been substituted by an alternative scene. The naturalness of the effect is subject to a number of defects which can detract from the desired illusion. In early equipment, which used a fast switching speed, noise on the derived keying signal gave rise to a disturbing edge jitter at the transition between the two signals. This problem is alleviated by reducing the gain in the keying channel and performing a fast cross-fade, rather than switching between the two sources. Unfortunately, this method results in some of the saturated keying colour appearing at the keying transition, giving rise to a coloured 'fringe' around the foreground subject; and further processing is necessary to suppress this undesirable effect. Nevertheless, with careful control of lighting conditions and foregound material colours, it is possible with modern keyers to obtain results which are virtually undetectable even if the foreground

includes smoke, glasses or other transparent objects.

With improvements in the technical performance of chromakeyers there has come a corresponding expansion in the range of applications of the technique. Hitherto, these have been restricted to the fields of news broadcasting, light entertainment, science fiction and advertising, where the technique is used as a special effect and is expected to be recognised as such. In these cases, the impact of the resulting effect is considered more significantly than any deficiencies there might be in the technical quality of the key. However, so far as serious productions are concerned, the use of chromakey to provide substitute background scenery must be undetectable by the viewer if the dramatic performance is not to be degraded. It is debatable whether present equipment meets this criterion, but the potential savings in scenery costs made possible by using chromakey-generated backgrounds are such as to make further improvements likely; and it is probable that chromakey techniques will assume increased importance in future television productions.

Consideration of digital television studio sampling standards must therefore include the question of chromakey performance. This is not only because of its increasing importance, but also because it is representative of the sort of complicated chrominance signal processing that will be undertaken in future digital television studios. The sampling in time and the quantising in amplitude of an analogue signal that are necessary for digital processing also introduce the corresponding distortions of aliasing and quantisation noise, which cannot later be removed. To see what effect these will have on digital chromakey equipment, a research project was initiated to investigate the relationships between sampling parameters, ease of implementation and subjective performance.

The Effect of Bandwidth Reduction and Noise on Chromakey Quality

In any typical chromakey production the main subject or subjects of the picture stand in front of a background of highly saturated colour. A signal is derived from the camera outputs to select a different background signal when the saturated colour is detected. Essentially, early chromakey designs switched from one picture to the other when the colour in question exceeded a preset limit. The equipment has since increased in sophistication such that the amount of new background added is now a fairly linear function of the amplitude of the designated colour present, or rather a function of a signal derived by a matrix from the three primary colours, so that the effects of luminance variations are cancelled. The generating of this keying signal means that the effect of colour spillage onto the foreground subjects can be reduced by subtracting the correct proportion of the keying signal. For complete removal of spillage it is essential that the wanted foregound subjects contain none of the keying colour. This is because the equipment cannot, for example, recognise the difference between blue eyes and a small amount of blue spillage onto blonde hair. The quality of a picture generated by use of chromakey depends critically on the accuracy with which the keying signal represents changes in colour of the original scene.

Production of the Keying Signal

The method of generating the keying signal has changed during the history of chromakey. Early equipment used a B–Y signal which eliminates the problem of keying on whites which would occur if B alone were used; but it suffers from the problem of keying from Magenta and Cyan, as well as Blue. A more complex matrix of the form: K (keying signal) = the greater of $(B-G)$, $(B-R)$ within the limits $0 \leqslant K \leqslant 1$, is currently popular. This matrix defines a range of colours between Magenta and Cyan as background. It is interesting to note that, because subtraction of colour signals is taking place, luminance changes take no part in the generating of the keying signal. The signal K cannot, however, be used directly to drive a high speed video fader to select foreground or background signals, since K is only unity when $B = 1$, $R = 0$, $G = 0$. This condition does not occur in signals from studio sources. With high quality background colouring and skilful lighting, it is possible to achieve levels of the order of $B = 0.7$, $R = 0.2$, $G = 0.35$. The maximum level of K will be 0.35, and a gain and clip function must be imposed to ensure complete selection of the new background signals. A minimum gain of $\times 3$ is required in this case, and it is found in practice that the elimination of shading and noise requires a gain of around $\times 5$. With a less pure background paint or cloth, and with lighting with noticeable shading or shadows, a gain of $\times 10$ might be required. This means that noise in the keying signal is increased by 14–20 dB before it is used to process the final picture.

The quality degradation due to noise in the foreground-background boundary region can be examined in terms of the fraction of the boundary region which is in error. We shall define as

objectionable any amplitude error in the keying signal of more than 10%. Examination of the keying signal in terms of its error function for a given signal-to-noise (snr) ratio will indicate the fraction in error. Ignoring aperture correction noise, which tends to be added equally in all three channels and so will be removed by the keying signal matrix, a studio camera has an unweighted snr of -45 dB, the noise exhibiting a flat frequency spectrum. The snr of K will be 14–20 dB worse, i.e. -25 to -31 dB. This is r.m.s. noise to peak-to-peak signal; so, for a lv signal with a 31 dB snr, the noise has an r.m.s. value of 28 mV. The probability of exceeding 100 mV is:

$$\left(1-\operatorname{erf}\frac{100}{28}\right) = 1-0.999644 = 0.0346\%$$

For 25 dB snr the corresponding figure is 7.42%.

If then, a gain of $\times 10$ is required for K, 7% of the transition region will be in error by more than 10%. This can be reduced only by filtering K to reduce the noise. Filtering to 3.5 MHz is required to reduce the fraction to 0.5% which is an acceptable figure.

The Effect of Bandwidth Reduction

For full resolution in the transition region the keying signal must have full system bandwidth. If bandwidth reduction is achieved by filtering K, then a softening of the transition region occurs. This is acceptable when keying large objects but causes problems when fine detail is being keyed, because the detail disappears. If K is generated from bandwidth-reduced signals a further problem arises; namely, because as an amplify-and-clip function is employed, the slowing down of transition edges by filtering can alter the timing of the keying signal. With a gain of $\times 5$ in K, the maximum error will be $T/4$; and, with a gain of $\times 10$ it will be $3\ T/8$ (where T is the risetime of the filtered edge).

TIMING ERROR AS A FUNCTION OF SYSTEM BANDWIDTH

BANDWIDTH (6-POLE ELLIPTIC FUNCTION FILTERS USED)	MAXIMUM TIMING ERROR (ns) GAIN = 5	GAIN = 10
1 MHz	113	170
2 MHz	57	85
3 MHz	38	57
4 MHz	28	43
5 MHz	23	34
6 MHz	19	28

A 34 ns error gives a 0.25 mm displacement on a 20-in diagonal screen. For an error of this magnitude or less, K must be generated from signals of at least a 4 MHz bandwidth.

From these considerations it can be seen that, for satisfactory chromakey in a realistic production, it is necessary to use source signals with at least 4 MHz bandwidth to derive the keying signal. The snr of the keying signal must be at least 27 dB; which means that, for studio camera sources, the key processing gain must not exceed 18 dB.

Implications for Digital Processing

The signal degradation introduced by a digital codec can be expressed in terms of three forms of distortion. These are:

(i) reduction in system bandwidth;
(ii) introduction of alias noise; and
(iii) introduction of quantising noise.

For a given sampling rate the first two are interdependent, greater bandwidth giving worse aliasing. For a given number range covered by the peak-to-peak video signal, the quantising noise is $q^2/12$, where q is the magnitude of one quantum.

The spectrum of this noise, and hence its visibility, is signal dependent. For a system with 4 MHz of chrominance bandwidth using 8-bit p.c.m. coding, the studio snr (46 dB) will not be degraded by the quantisation noise of the digital codec (58 dB). Aliasing, however, is less well defined. The keying signal tends to contain high frequency information, as much of the transition region consists of sharp edges. This is exacerbated by non-linear processing, which generates harmonics which appear outside the permitted frequency band of $0 \leqslant f \leqslant f_s/2$—($f_s$ being the sampling frequency). If the processing is performed at a sufficiently high clock-rate this effect can be reduced; but aliasing introduced in the initial sampling process cannot be removed without loss of high frequency information.

The worst type of signal for generating alias components will be a comb structure with spatial frequency of 400 lines per picture width. To prevent alias visibility for this type of signal the pre-sampling filter must have 35 dB attenuation at $f_s/2$.

In a realistic production the type of foreground most likely to cause aliasing problems is one with thin vertical or near-vertical objects. With an orthogonal sampling pattern the alias distortion is displayed as a step effect on near-vertical edges. The steps remain stationary as the object moves, and are very obviously not part of the original scene. Thin objects in the

foreground give rise to signals of the form of 1 T pulses. The amplitude spectrum of such a pulse is zero at 10 MHz, passing through −17.9 dB at 6.75 MHz—i.e. $f_s/2$ for a sampling rate (f_s) of 13.5 MHz. To reduce aliasing to an acceptable level for such a signal, the filter requirement is reduced to approximately 17 dB at $f_s/2$. If a clock of lower frequency is used, the filter requirement becomes more critical.

An Experimental Digital Chromakeyer

Early discussions within the European Broadcasting Union gave rise to a preliminary proposal for a digital sampling standard of 12 MHz for the luminance and 4 MHz for each of the colour difference signals ($768 f_h$ and $256 f_h$ respectively). Several organisations undertook to investigate various aspects of the suitability of this proposal, and the system which the IBA developed is shown in Fig. 1. An 8×8 three-channel assignment matrix, operating in parallel mode, provides a full range of switching options between the various picture sources and feeds, and also includes variable digital delays so that the complete system timing can readily be adjusted. The chromakey unit has three feeds:—foreground, background and keying source. This last, in normal operation, is the same as the foreground. Various inputs may be assigned to these feeds including two YUV inputs (camera or slide-scanner), internally-generated colour fields or the output of a digital VTR, so that 'downstream' chromakeying from a recorded foreground subject can be demonstrated. The complete system, which contains representative elements of an all-component digital studio, is controlled by a microprocessor operating under instructions from a remote operating console.

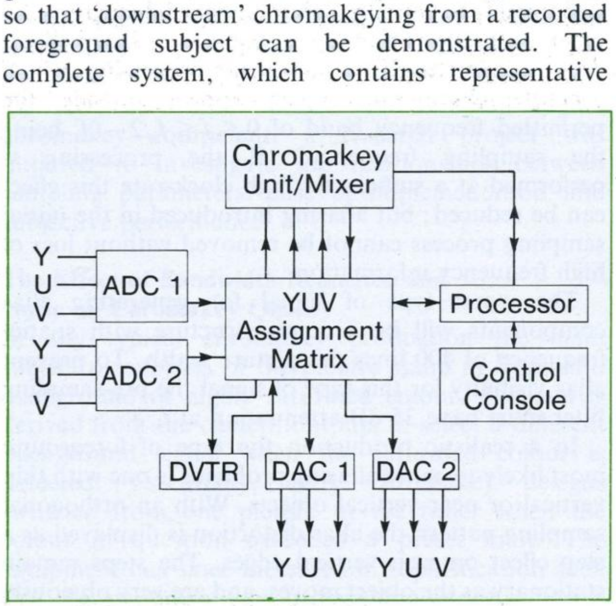


Fig. 1. (12 : 4 : 4) sampling system.

Key Signal Generation

To ensure accurate selection of keying colour, samples of the keying area of the foreground are taken direct from the foreground signal. The operator achieves this in practice by superimposing a pair of cross-wires over a representative portion of the keying area. Upon an input command, U and V samples from the foreground source are latched into the control computer and used as background co-ordinates U_0 and V_0. During keying operations each pair of chrominance samples is operated upon to generate a measure of its distance in the U, V plane from the background point U_0, V_0. The first operation is a co-ordinate translation and rotation giving new co-ordinate axes centred on the background point and aligned along the hue and saturation directions, according to the following equations:

$$u' = (u-u_0)\cos\phi + (v-v_0)\sin\phi$$

$$v' = (v-v_0)\cos\phi - (u-u_0)\sin\phi$$

The keying signal is then formed:

$$K = a|u'| + b|v'|$$

The locus of constant K is a rhombus in the u–v plane, the shape of which is controlled by the gain factors a and b given above. Up to this stage all signal processing has been linear, and hence can be performed at the input frequency of 4 MHz.

This 'distance' function is then followed by an adjustable threshold and a variable gain and clip with the result that the keying signal as a function of u and v input signals is typically as shown in Fig. 3. For a range of values around the background point, the keying signal is zero, the size of this area being determined by the signal threshold. For values outside the larger area the keying signal is unity. Increasing the keying gain reduces the size of the transition region. Since this processing is non-linear, harmonics of the keying signal will be generated, giving rise to severe aliasing particularly if a high gain (and hence a large amount of clipping) is used. To alleviate these effects the signal is interpolated to 12 MHz before the threshold and gain are applied.

The eventual keying signal is then used to perform an electronic 'cross-fade' between the foreground and background according to the form:

$$Kf + (1-K)b$$

where f and b are the foreground and background signals respectively.

The duration of this cross-fade is therefore

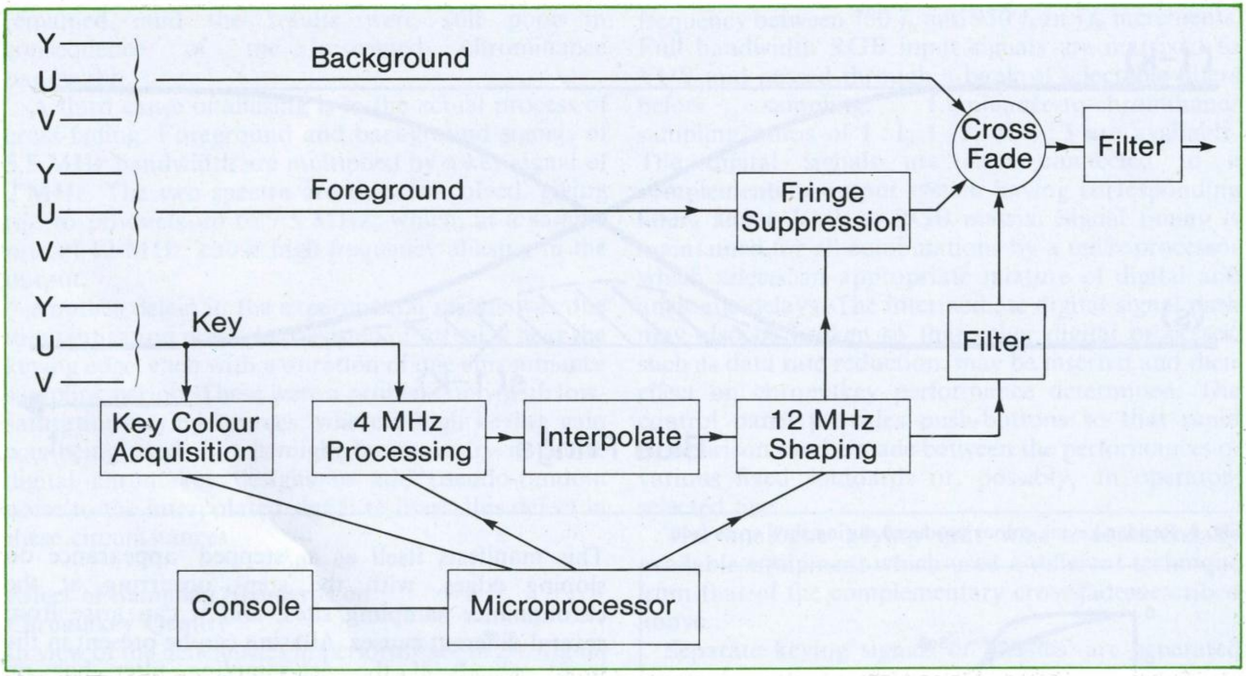


Fig. 2. Schematic—chromakey unit.

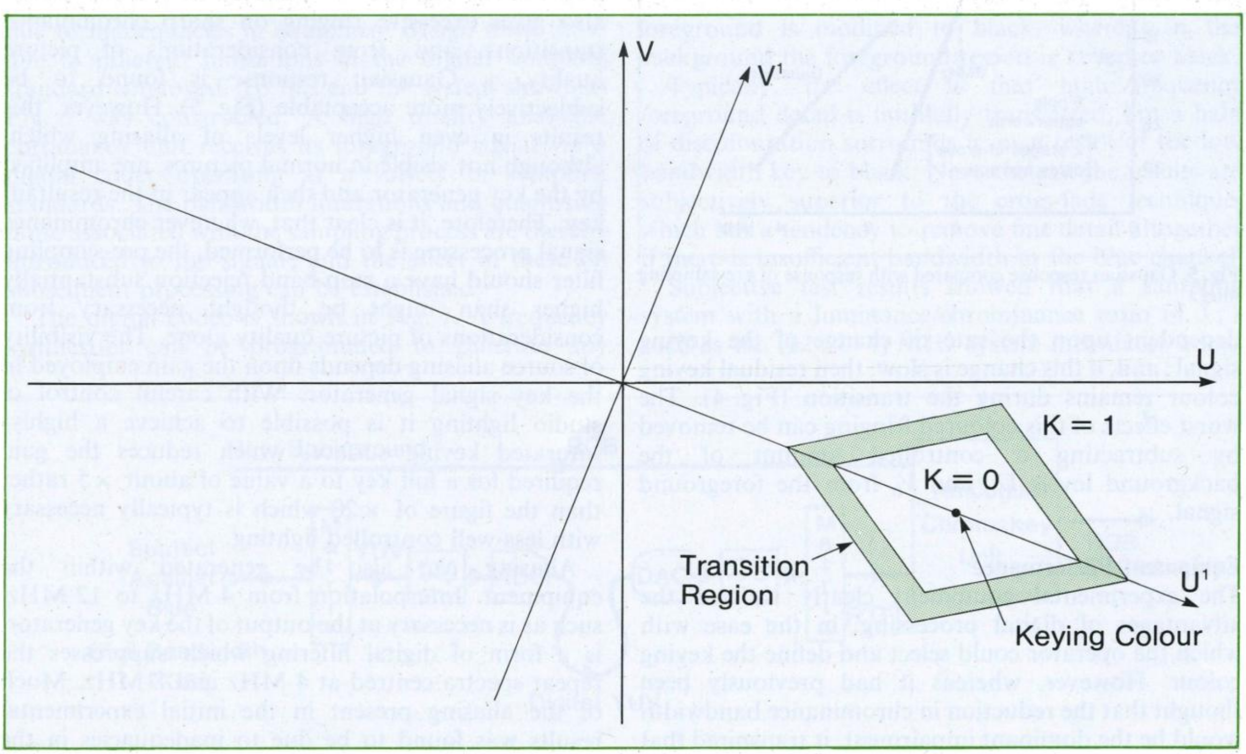


Fig. 3. Keying signal as a function of 'u' and 'v' input signals.

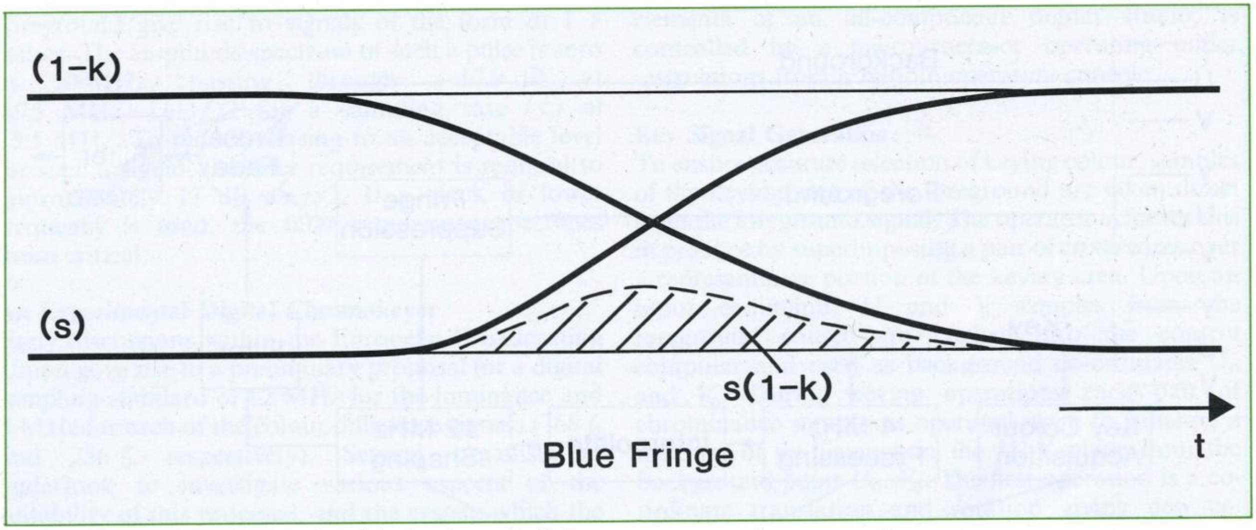


Fig. 4. Residue keying colour produced during slow cross-fade.

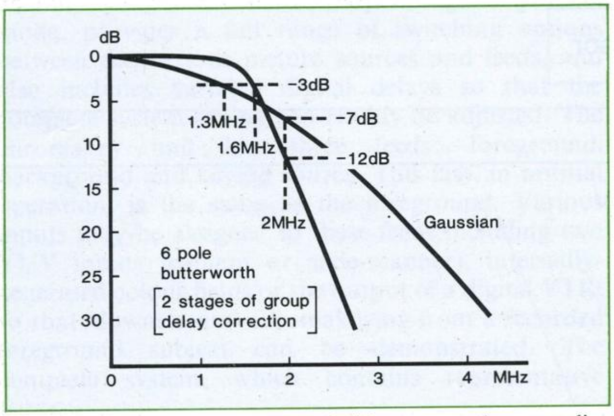


Fig. 5. Gaussian response compared with response of pre-sampling filter.

dependent upon the rate of change of the keying signal; and, if this change is slow, then residual keying colour remains during the transition (Fig. 4). The worst effects of this coloured fringing can be removed by subtracting a controlled amount of the background levels U_0 and V_0 from the foreground signal.

Equipment Performance

The experimental equipment clearly showed the advantages of digital processing, in the ease with which the operator could select and define the keying colour. However, whereas it had previously been thought that the reduction in chrominance bandwidth would be the dominant impairment, it transpired that chrominance aliasing was much more objectionable. This manifests itself as a 'stepped' appearance on sloping edges, with the steps occurring at the chrominance sampling sites, and it can arise from several different causes. Aliasing can be present in the input signals if the pre-sampling filter has an insufficiently fast cut-off. Unfortunately, a fast cut-off also gives excessive ringing on sharp chrominance transitions, and, from considerations of picture quality, a Gaussian response is found to be subjectively more acceptable (Fig. 5). However, this results in even higher levels of aliasing which, although not visible in normal pictures, are amplified by the key generator and then appear in the resultant key. Therefore, it is clear that, wherever chrominance signal processing is to be performed, the pre-sampling filter should have a stop-band rejection substantially higher than might be thought necessary from considerations of picture quality alone. The visibility of source aliasing depends upon the gain employed in the key signal generator. With careful control of studio lighting it is possible to achieve a highly-saturated keying surface, which reduces the gain required for a full key to a value of about $\times 5$ rather than the figure of $\times 20$ which is typically necessary with less-well controlled lighting.

Aliasing can also be generated within the equipment. Interpolation from 4 MHz to 12 MHz, such as is necessary at the output of the key generator, is a form of digital filtering which suppresses the repeat spectra centred at 4 MHz and 8 MHz. Much of the aliasing present in the initial experimental results was found to be due to inadequacies in the interpolator used. However, some traces of aliasing

remained, and the results were still poor in consequence of the restricted chrominance bandwidth.

A third cause of aliasing is in the actual process of cross-fading. Foreground and background signals of 5.5 MHz bandwidth are multiplied by a key signal of 2 MHz. The two spectra are thus convolved, giving rise to products up to 7.5 MHz; which, at a sample rate of 12 MHz, cause high frequency aliasing in the output.

Another defect in the experimental results was due to quantisation noise. It consisted of 'streaks' near the keying edge, each with a duration of one chrominance sampling period. These were a problem only with low-saturation keying sources, where a high keying gain was being used; but it might be necessary in future digital chromakey designs to add pseudo-random noise to the interpolated signal to mask this defect in these circumstances.

Effect of Sampling Standards on Chromakey Quality

In view of the deficiencies in performance of the digital chromakey equipment, it was decided to undertake a separate investigation which would dissociate defects due to inadequacies in equipment design from those due to inherent limitations in the digital sampling standard employed. To this end the system shown in Fig. 6 was constructed. A high quality analogue chromakey unit receives its foreground signal via a digital codec operating at a variety of sampling standards. The bandwidth limitations and quantising noise associated with the sampling process are thereby introduced to the signal; and the effect of these on subsequent processing can be established.

The digital codec is shown in Fig. 7. A frequency synthesiser can be programmed to generate any frequency between 750 f_h and 930 f_h in $\frac{1}{2}f_h$ increments. Full bandwidth RGB input signals are matrixed to YUV and passed through a bank of selectable filters before sampling. Luminance-to-chrominance sampling ratios of 1 : 1, 1 : 2, or 1 : 3 are available. The digital signals are then connected to a complementary output system having corresponding filters and a YUV-to-RGB matrix. Signal timing is maintained for all combinations by a microprocessor which selects an appropriate mixture of digital and analogue delays. The intermediate digital signal path may also be broken so that other digital processes, such as data rate reduction, may be inserted and their effect on chromakey performance determined. The control panel provides push-buttons so that rapid comparison can be made between the performances of various fixed standards or, possibly, an operator-selected one.

The analogue keying unit was a commercially available equipment which used a different technique from that of the complementary cross-fade described above.

Separate keying signals or 'mattes' are generated for the foreground and background signals which are then added together linearly. The keying colour in the foreground is modified to black, whereas in the background the foreground region is keyed to black.

Typically, the effect is that high frequency foreground detail is faithfully transferred, but a halo of discolouration surrounds it, as a result of the low bandwidth key to black. Nevertheless, the results are subjectively superior to the cross-fade technique, which has a tendency to remove fine detail altogether if there is insufficient bandwidth in the blue channel.

Subjective test results showed that a sampling system with a luminance/chrominance ratio of 3 : 1 such as the (12 : 4 : 4) MHz system introduces a loss

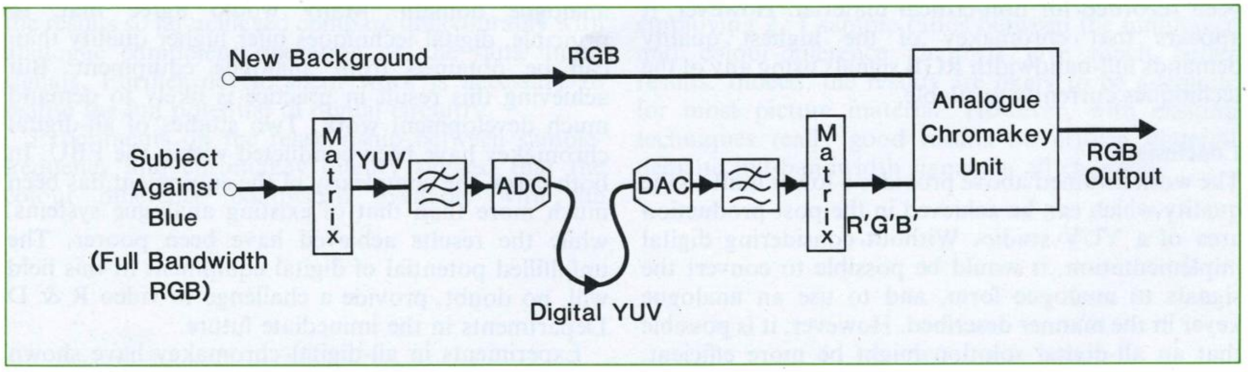


Fig. 6. Block diagram of equipment designed to investigate inherent limitations in digital sampling standard used.

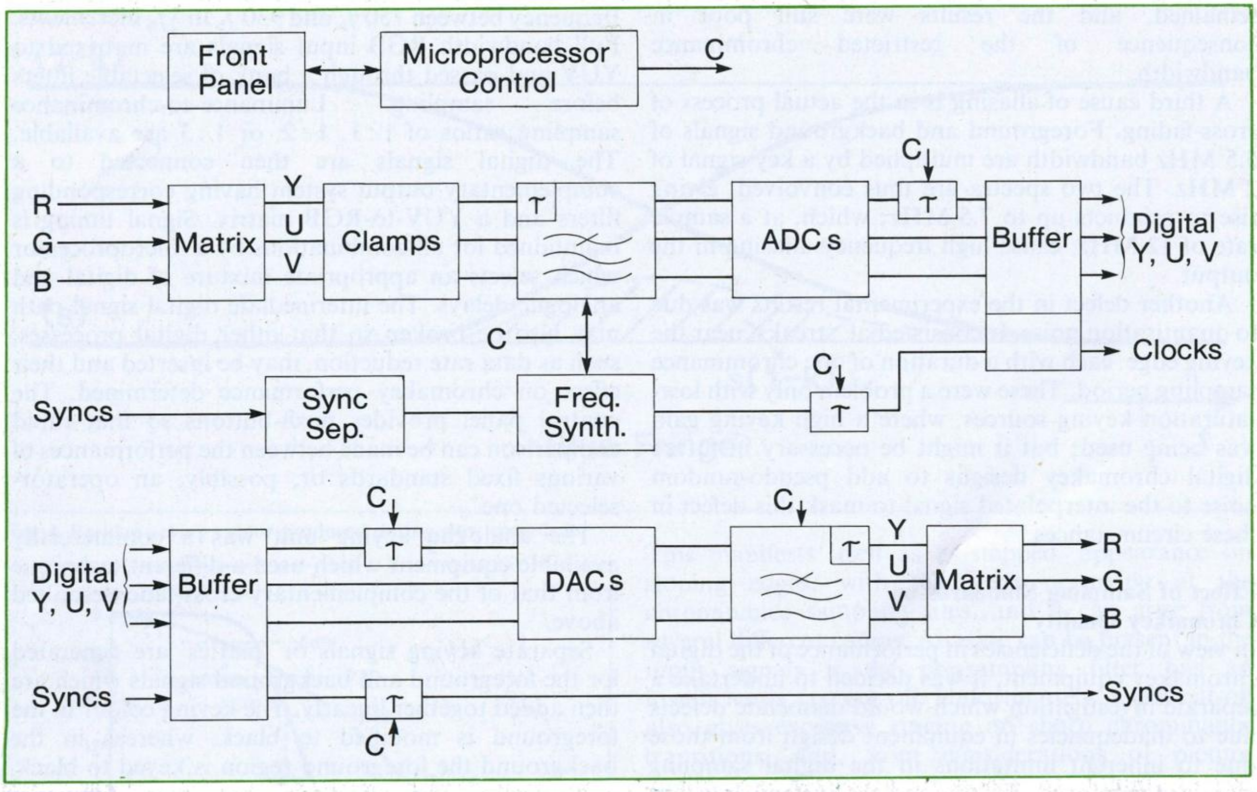


Fig. 7. Digital codec.

of quality amounting to 0.95 of a picture grade. This might be attributable to a lack of resolution in the colour difference channel. All of the 2 : 1 ratio systems tested (in the range 12 : 6 : 6 MHz to 14 : 7 : 7 MHz sampling) introduce a quality loss in the region of 0.5 of a grade, the improvement through this range amounting to about 0.1 of a grade. It is true to say that the choice of test material would certainly affect the result, and that lower impairments would have been recorded for non-critical material. However, it appears that chromakey of the highest quality demands full-bandwidth RGB signals using any of the techniques currently available.

Conclusions

The work outlined above provides a lower limit to the quality which can be achieved in the post-production area of a YUV studio. Without considering digital implementation, it would be possible to convert the signals to analogue form, and to use an analogue keyer in the manner described. However, it is possible that an all-digital solution might be more efficient, and might produce results of a higher quality. Information theory suggests that the digital technique should be capable of providing results at least comparable with those of analogue equipment when provided with the same input signals. It has been found, however, that the signal processing involved in achieving this result can be quite complex. Conversely, the increased flexibility of digital techniques in the storage and manipulation of signals might offer possibilities which are not available in the analogue domain. Many would agree that, in principle, digital techniques offer higher quality than can be obtained from analogue equipment. But achieving this result in practice is likely to demand much development work. Two studies of all-digital chromakey have been conducted within the EBU. In both cases, the complexity of the equipment has been much more than that of existing analogue systems, while the results achieved have been poorer. The unfulfilled potential of digital equipment in this field will, no doubt, provide a challenge to video R & D Departments in the immediate future.

Experiments in all-digital chromakey have shown that it is very difficult to achieve good results when

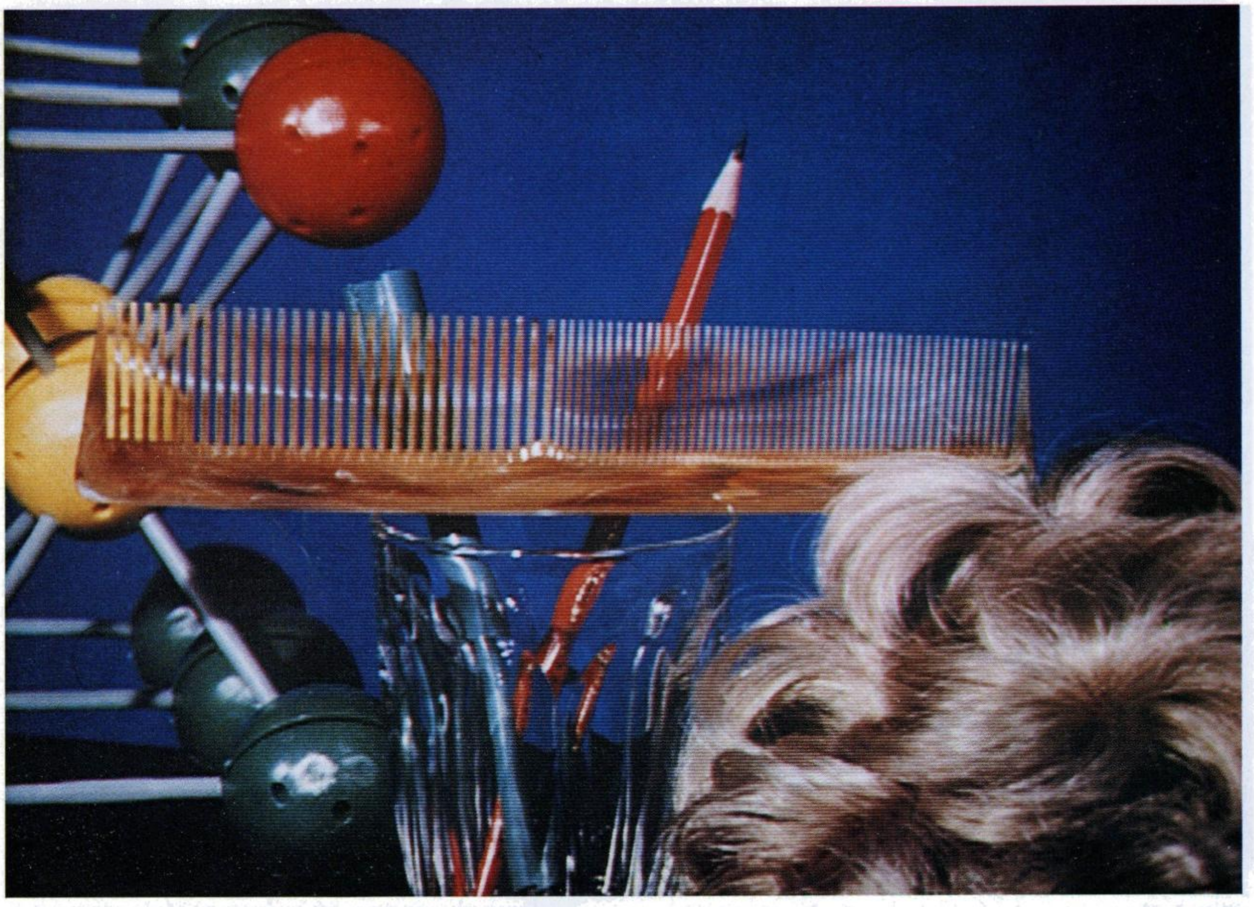


Plate 1. This shows the varied subject material used as sources for chromakey experiments.

using the 'cross-fade' technique alone, particularly when the bandwidths of the colour difference channels are restricted. Significant improvements can be made by employing the 'additive colour-matte' technique. However, the equipment design is very complex, and the results so far achieved compare unfavourably with analogue equipment operating on the same input signals. Further development work is necessary to realise the full potential of digital circuitry.

Experiments on the relationship between sample-frequency and chromakey quality suggest that the colour difference bandwidth available with the (12 : 4 : 4) MHz proposal are insufficient to provide high-quality results.

This result has been confirmed in analogue and digital experiments using both the 'cross-fade' and 'additive matte' methods. Sampling standards employing 2 : 1 sample ratios between the luminance and colour-difference channels give much better results. Indeed, the results are likely to be adequate for most picture material. However, with existing techniques really good results on critical material demand full-bandwidth signals in all channels.

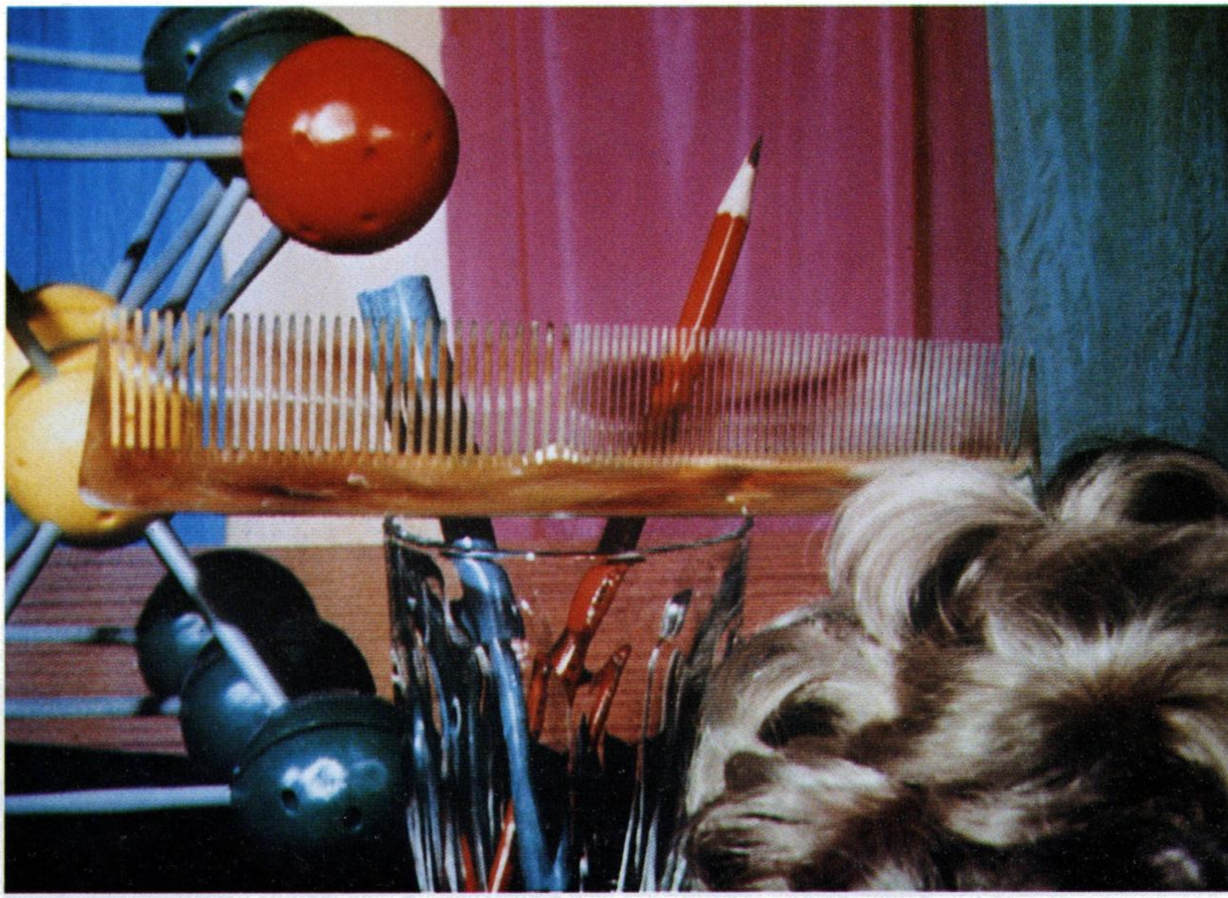


Plate 2. The keyed result obtained when the foreground signal is sampled at (14 : 14 : 14) MHz.

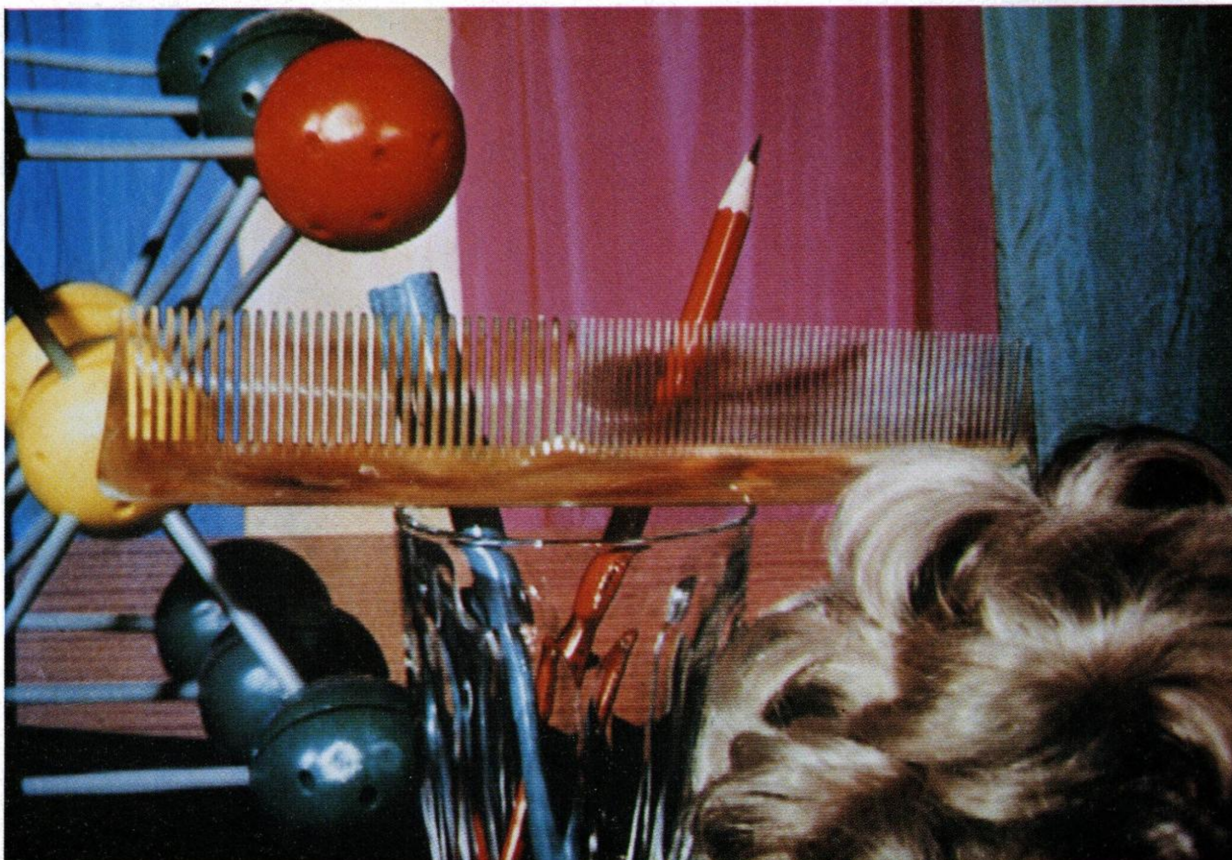


Plate 3. The result obtained with a sampling standard of (14 : 7 : 7) MHz. The reduced chrominance bandwidth gives rise to poorer keying in the region of the fine teeth of the comb near the pencil.

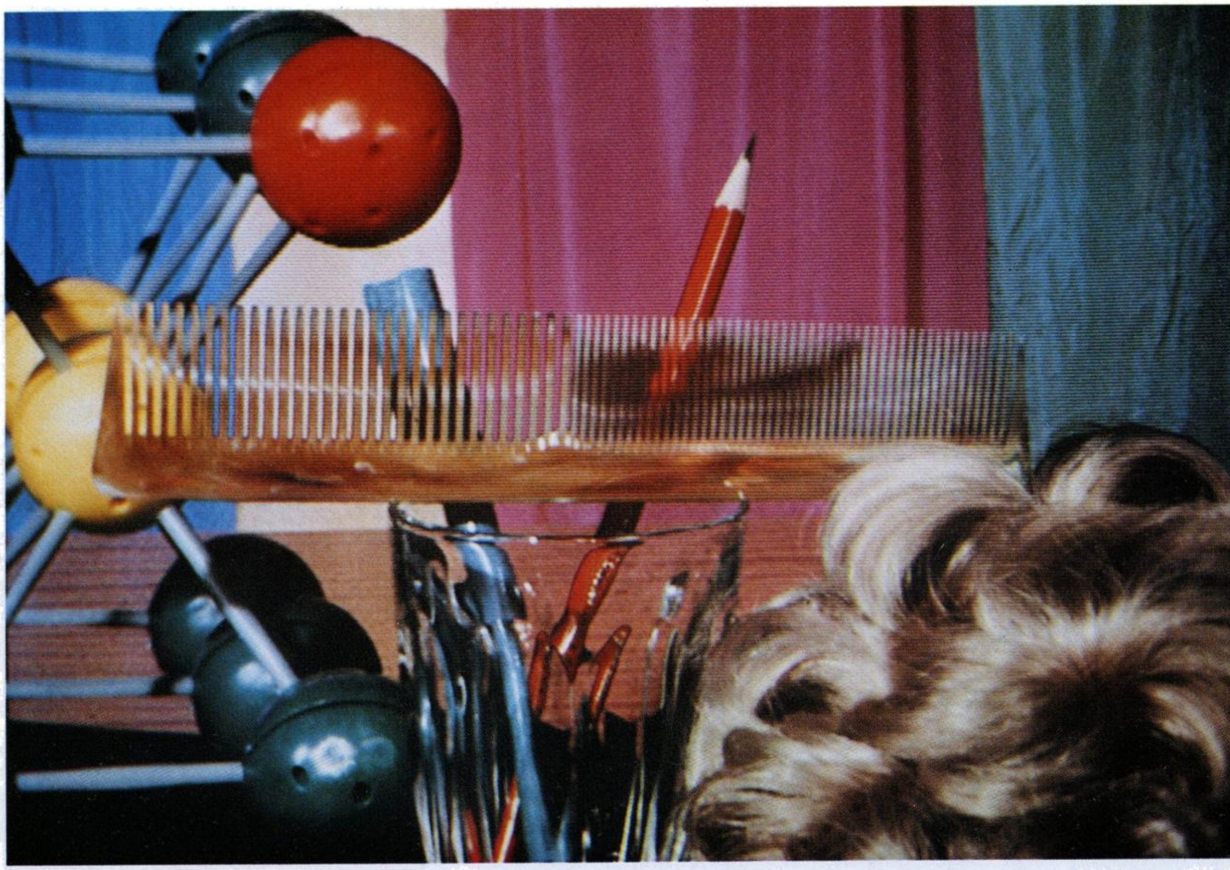


Plate 4. A sampling standard of (12 : 6 : 6) MHz. A further deterioration in the fine teeth of the comb can be seen.

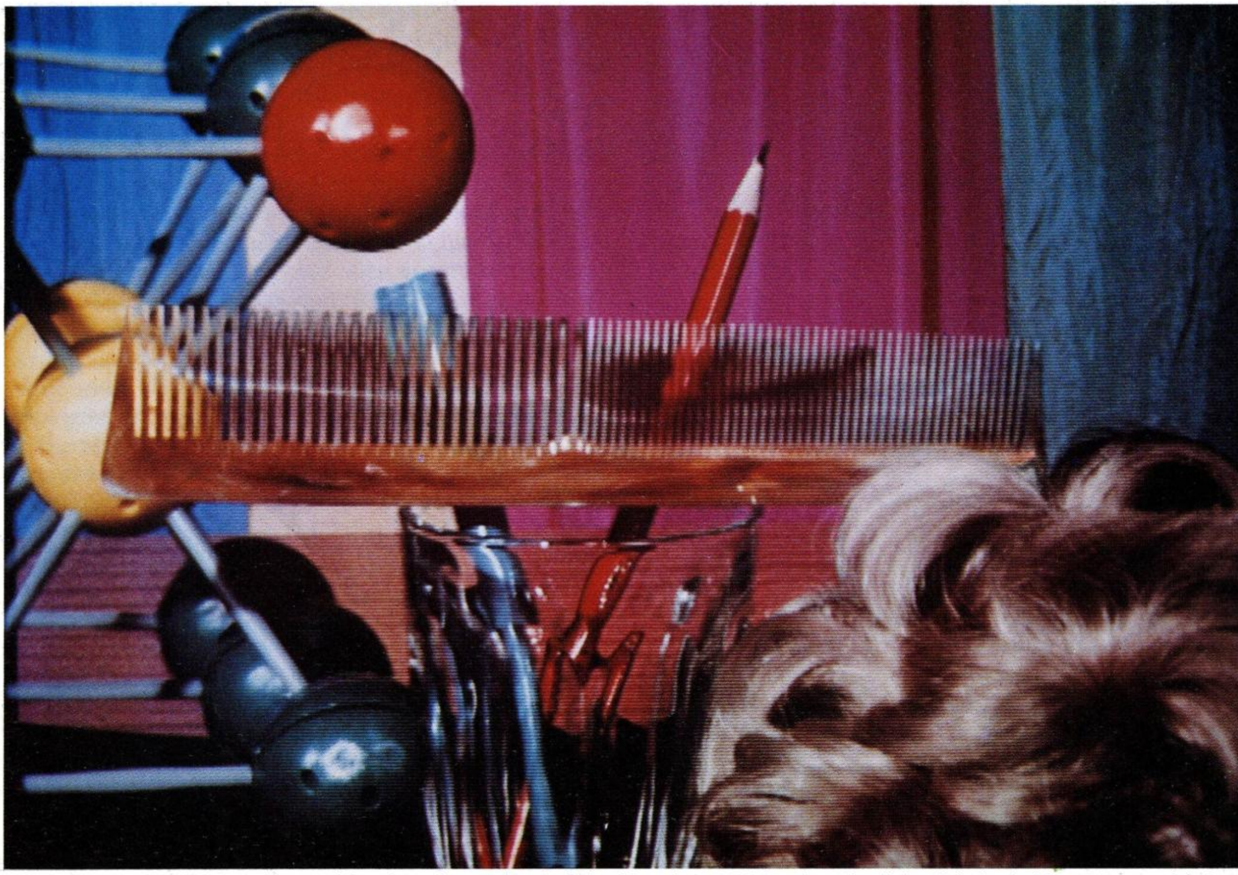


Plate 5. (12 : 4 : 4) MHz sampling. With a chrominance bandwidth of less than 2 MHz, virtually all detail in both sides of the comb is lost. A reduction in the saturation of the blue signal in these regions leads to a 'soft mix' between the foreground and background instead of a proper 'key'.

EDGAR WILSON,
B.Sc.(Eng.), ACGI.
A biographical note appears on page 12.

PETER CARMEN,
C.Eng., MIEE,
completed a six-year apprenticeship at the Royal Radar Establishment, Great Malvern, and subsequently worked on industrial television and military equipment with the Marconi Company. He joined the IBA's Experimental and Development Department in December 1973 and since then his work in digital signal processing has included designs for colour coding systems, very high speed serial arithmetic and participation in equipment design and development for coding and transmission of digital television signals via satellite.
He is married and lives in rural Wiltshire.

Bit-rate Reduction for 140 Mbit/s Links

by E. J. Wilson and P. R. Carmen

Synopsis
Recent advances in technology, particularly in the area of fibre optic cables, have brought forward the planned introduction of high capacity digital links.

The 140 Mbit/s link at the European standard fourth order multiplex level, may enable the economic *conversion* of the total network to digital operation—perhaps within a decade. For broadcasters a totally digital network of 140 Mbit/s links would allow us to contemplate the superior quality possible by transmission of digital television signals in component form. The text describes investigative work carried out to optimise the subjective quality of a 140 Mbit/s bit-rate reduced signal and gives details of the preferred method of signal processing which allows virtually impairment free link transmission.

Introduction

In the past, when digital television standards for studio use have been discussed, secondary issues such as the present bit-rate limits of digital videotape recorders and economic transmission via digital links have been allowed to cloud our view of the studio signal primary requirements. More recently by divorcing the studio signal, temporarily, from external influences, we have concentrated on the selection of a suitable standard for luminance and chrominance sampling frequencies which will cope simply with the present signal processing needs, such as for chromakey and electronic picture storage, and with foreseeable requirements, such as those for special effects.

By mid-1980 the IBA was concentrating attention on standards of which two common characteristics were that the luminance to chrominance sampling ratio should be 2 : 1 : 1 and that the three component signals should be orthogonally sampled using 8-bit quantisation. On this basis the effective serial bit-rate of a signal, sampled at even the lowest possible proposed frequency of 12 MHz, would be 192 Mbit/s. On leaving the studio centre the signal would require compression to less than 140 Mbit/s in order to utilise the fourth order multiplex PTT links at this data rate. In late 1980 the IBA began a 140 Mbit/s link simulation project in order to examine methods of bit-rate reduction for component TV signals. The primary purpose was to enable 140 Mbit/s digital link transmission of processed signals originated at any of the proposed sampling frequencies; but a strong secondary requirement was that emphasis should be given to retaining both picture quality and

'downstream' signal processing capability.

A third aim was that in meeting our other objectives we should use only methods which would be cheap and simple to implement in a final equipment.

Several of our ideas, described in detail later, were implemented in a flexible simulation equipment with which we were able to compare the various methods. As a result, one technique which gave excellent quality pictures and enabled very good 'downstream chromakey' was demonstrated in January 1981 to delegates from the EBU and SMPTE. In subsequent formal subjective testing this method of bit-rate reduction was judged to cause no significant impairment of the original studio standard signal.

The remainder of this chapter describes in detail this preferred method of bit-rate reduction together with the other methods which could be simulated on the demonstration equipment. One method of achieving a reduction to less than half the studio bit-rate was also included in the equipment. This technique which may prove useful for ENG applications is described in the earlier chapter 'An Extensible Family of Standards'.

What is 140 Mbit/s?

140 Mbit/s, or more precisely 139.264 Mbit/s, is the European standard fourth order multiplex digital transmission level. Its value results from the combining in lower order multiplexers of $30 \times 4 \times 4 \times 4$ telephone channels each of 64 kbit/s and the precise bit-rate takes account of the necessary housekeeping overheads incurred at each multiplex stage.

The present microwave radio relay network utilises the same channel spacing for analogue television as for telephony FDM (Frequency Division Multiplex) Hypergroups. We can, therefore, equate a PAL System I video signal with 960 telephone channels in SHF bandwidth requirement.

A 140 Mbit/s link normally carries 1,920 telephone channels; and we might initially consider this to be the equivalent of two digitised PAL signals. With a component television signal, which has not suffered the degradation inherent in PAL bandwidth sharing, it might be considered economic, therefore, to accept the bit-rate penalty of utilising a complete 140 Mbit/s link per television channel. Within this digital link an allocation of some bit-rate for audio, asynchronous multiplexing and other system housekeeping must be made. Assuming 5 Mbit/s overhead for this, we are left with about 134 Mbit/s for the coded video signal; and it is from this value that the effective number of bits per sample for various possible sampling standards, as shown in Table 1, have been computed.

One common element of bit-rate reduction assumed for any of these preferred standards, which sample luminance and chrominance in a 2 : 1 : 1 ratio, is line blanking removal. This gives a theoretical reduction of 18% bit-rate at the expense of some buffer storage and digital 'line sync' housekeeping overhead, and causes no impairment to the active picture area. A further 8% reduction in bit-rate could be achieved by removing the field blanking interval, but the storage required for this is considered to be outside our requirement for simplicity and so is not further considered.

Accepting, therefore, that 134 Mbit/s of link capacity must carry the line blanking removed signal, the effective average word length remaining varies as shown from 5.79 to 6.87 for the different possible sampling standards. Several methods of achieving these orders of effective word lengths are described below.

Bit-rate Reduction Methods

There are two main ways in which the effective average word length can be reduced from its initial value of 8-bits. The first, which involves discarding

TABLE 1: POSSIBLE COMPONENT TV DIGITAL STANDARDS

STANDARD (Y : U : V) MHZ	NO. OF LUMINANCE SAMPLES/ LINE	SOURCE DATA RATE (MBIT/S)	AFTER LINE BLANKING REMOVAL (MBIT/S)	AVERAGE WORD LENGTH REMAINING (BITS)
12 : 6 : 6	768	192	156	6.87
13 : 6.5 : 6.5	832	208	169	6.34
13.5 : 6.75 : 6.75	864	216	175.5	6.11
14 : 7 : 7	912	228	185.25	5.79

alternate samples, is known as sub-Nyquist sampling since it involves a final sample rate less than the Nyquist limit of twice the maximum frequency of the baseband signal. In practice, advantage is taken of the two-dimensional sampling nature of the digitised television signal to maintain horizontal and vertical spatial frequency response at the expense of some loss of resolution and introduction of aliasing in diagonal frequencies. The second technique involves differential pulse code modulation where the redundancy in a typical television signal (the 'sameness' of samples within small areas of the picture) is removed by transmitting only difference information. The difference information is coded in compressed non-linear form, and this gives rise to increased quantisation noise in areas of the picture with little 'sameness'.

The modes of operation simulated by our equipment are variations and combinations of these two techniques.

MODE 1—HYBRID DIFFERENTIAL PULSE CODE MODULATION (HDPCM), 8–5–5–5. Every fourth word in a line quincunx structure remains as an unmodified 8-bit pulse code modulated sample. The pattern of these words is then used as the basis of two-dimensional predictor/interpolators for the intervening words which are non-linearly coded 5-bit differences. This results in an average word length of 5.75-bits.

MODE 2—HDPCM 8–$4\frac{1}{2}$–$4\frac{1}{2}$. This is similar to Mode 1 except that every third word is unmodified, the other two in each group of three being coded as differences of $4\frac{1}{2}$-bits. The average word length is 5.67-bits.

MODE 3—HDPCM 8–4. Alternate words in a line quincunx structure are either 8-bit pcm samples or 4-bit differences giving an average word length of 6-bits.

MODE 4—SUB-NYQUIST CODING 8–0. An average word length of 4-bits is achieved by discarding every second sample for transmission. Programmable two-dimensional transversal filters are used to reduce diagonal resolution, and aliasing, prior to sample omitting and reconstructing.

MODE 5—SUB-NYQUIST CODING 8–4. This is similar to Mode 4 but the alternate 8-bit samples are supplemented with 4-bit difference words to remove alias products remaining in spite of the two-dimensional filtering. These difference words increase the average word length to 6-bits. In this Mode the residual aliasing present in Mode 4 is replaced by an increase in quantising noise.

MODE 6—SUB-NYQUIST WITH FRAME RESET 8–0. The sub-Nyquist coding of Mode 4 suffers from a repetitive sequence of four field types which could make for difficulty in videotape editing and electronic picture storage. Similar picture quality without this disadvantage can be obtained by resetting the sampling sequence once every two fields. The average word length of 4-bits is unaffected.

In demonstration and subsequent quantative subjective testing Mode 1 gave outstanding results. In addition, at 5.75-bits/sample it is suitable for use with even the highest proposed sampling standard of (14 : 7 : 7) MHz. The method is therefore fully described below with further discussions of Modes 2 and 3 confined to an appendix. The sub-Nyquist techniques of Modes 4, 5 and 6 are described in detail in the appendix to the chapter on extensible families.

The Two-dimensional Hybrid Differential Pulse Code Modulation Mode 1

This method can be applied to any orthogonally sampled digital signal. In the demonstrated equipment the sampling frequency was 14.25 MHz for luminance (912 samples per television line) and 7.125 MHz for each of the chrominance signals (456 samples per line).

Figure 1 illustrates the approximate spatial pattern produced by the luminance sample sites of three television lines.

One in every four of these source samples is labelled as a PCM (P) sample in a line quincunx structure. The remaining samples are labelled as DPCM (D) samples (Fig. 2). In the 140 Mbit/s transmission signal the 'P' samples remain as 8-bit unmodified numbers while the 'D' samples are compressed to 5-bits each. The

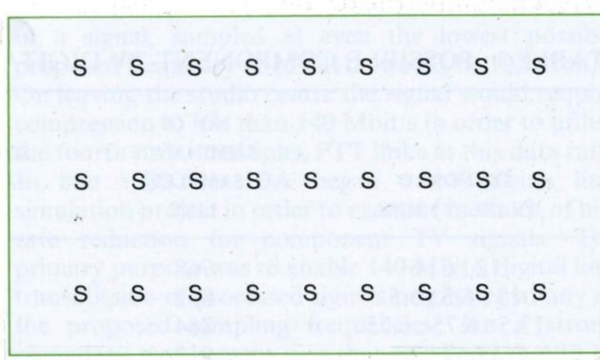


Fig. 1. Luminance sample sites.

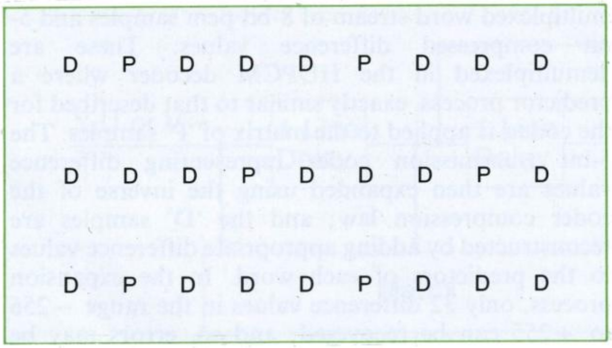
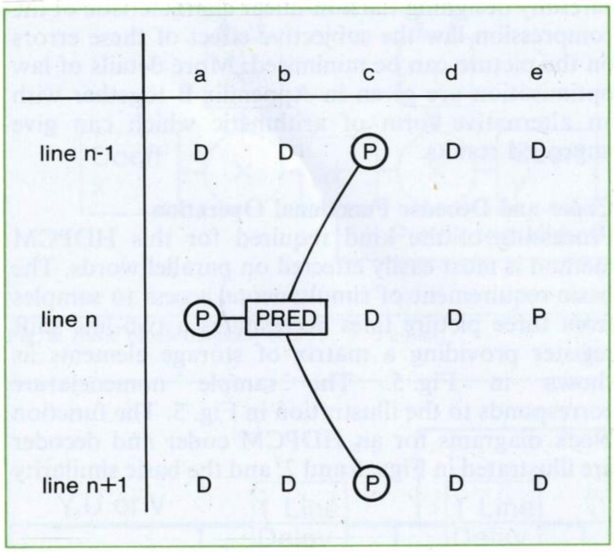
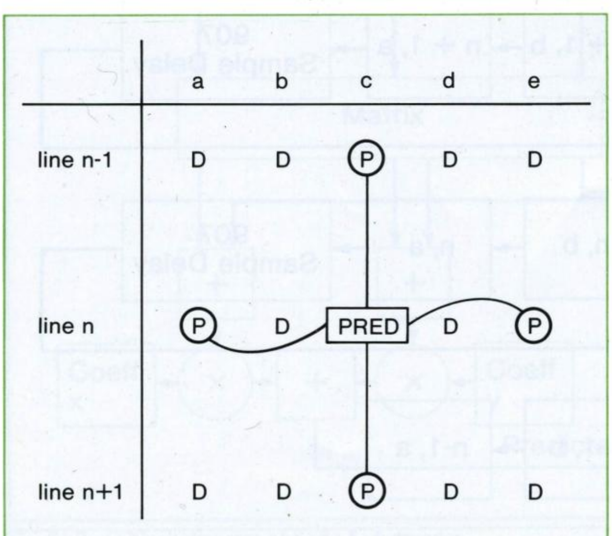


Fig. 2. Line quincunx PCM (P) sample structure.





first stage in this process is the calculation of a 'prediction' for each of the 'D' samples. This prediction is made by calculating the weighted average of several of the 'P' samples in the near vicinity of the 'D' sample in question. This process is illustrated in Fig. 3 for the three types of predicted value where the calculations are made according to the formulae:

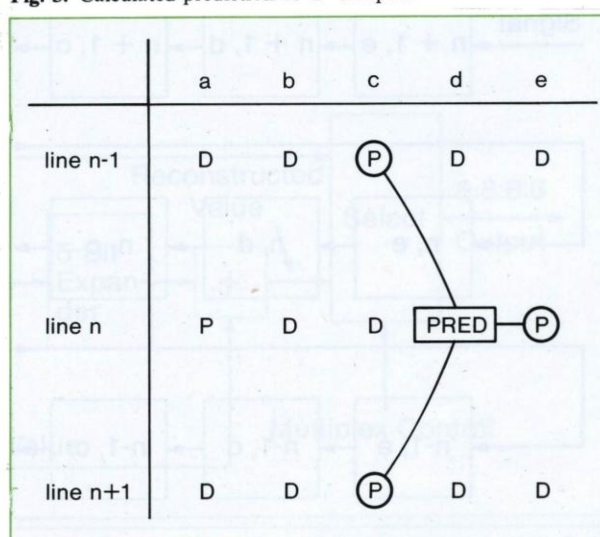
$$\mathrm{PRED}(n, b) = xP(n, a) + y[P(n-1, c) + P(n+1, c)]$$
$$\mathrm{PRED}(n, c) = x[P(n, a+P(n, e)] + y[P(n-1c) + P(n+1, c)]$$
$$\mathrm{PRED}(n, d) = xP(n, e) + y[P(n-1, c) + P(n+1, c)]$$

The values of the coefficients x and y differ depending upon the predictor type and whether the samples are, of luminance or chrominance. The coefficients used in practice, which were rounded to simple fractions for ease of implementation, are tabulated below.

TABLE 2: PREDICTOR COEFFICIENTS

	PREDICTOR	X	Y
Luminance	(n.b)	$\frac{1}{2}$	$\frac{1}{4}$
	(n.c)	$\frac{1}{4}$	$\frac{1}{4}$
	(n.d)	$\frac{1}{2}$	$\frac{1}{4}$
Chrominance	(n.b)	$\frac{3}{8}$	$\frac{5}{16}$
	(n.c)	$\frac{3}{16}$	$\frac{5}{16}$
	(n.d)	$\frac{3}{8}$	$\frac{5}{16}$

Fig. 3. Calculated prediction of 'P' samples.



For optimum results at sampling rates other than (14 : 7 : 7) MHz it might prove necessary to alter these values.

Once the predictor for each 'D' sample position has been obtained, difference values formed by subtracting the predictor from the real 'D' sample are obtained. If 2's complement arithmetic is used in this process, difference values from −256 to +255 must be compressed for transmission into 32 levels. This is done by using a non-linear law such as illustrated in Fig. 4.

The output of the HDPCM coder is then a multiplexed word stream of 8-bit pcm samples and 5-bit compressed difference values. These are demultiplexed in the HDPCM decoder where a predictor process, exactly similar to that described for the coder, is applied to the matrix of 'P' samples. The 5-bit transmission codes representing difference values are then expanded using the inverse of the coder compression law; and the 'D' samples are reconstructed by adding appropriate difference values to the predictors of each word. In the expansion process, only 32 difference values in the range −256 to +255 can be recovered; and so, errors may be included in the reconstructed 'D' samples. By carefully designing the non-linear characteristic of the compression law the subjective effect of these errors on the picture can be minimised. More details of law optimisation are given in Appendix B together with an alternative form of arithmetic which can give improved results.

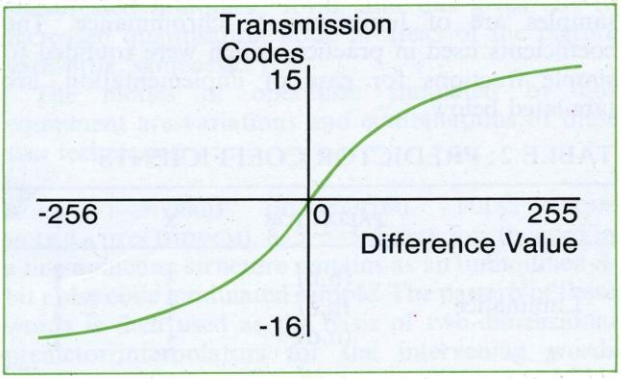


Fig. 4. 2's complement companding law.

Coder and Decoder Functional Operation

Processing of the kind required for this HDPCM method is most easily effected on parallel words. The basic requirement of simultaneous access to samples from three picture lines necessitates a two-line shift register providing a matrix of storage elements as shown in Fig. 5. The sample nomenclature corresponds to the illustration in Fig. 3. The function block diagrams for an HDPCM coder and decoder are illustrated in Figs. 6 and 7, and the basic similarity

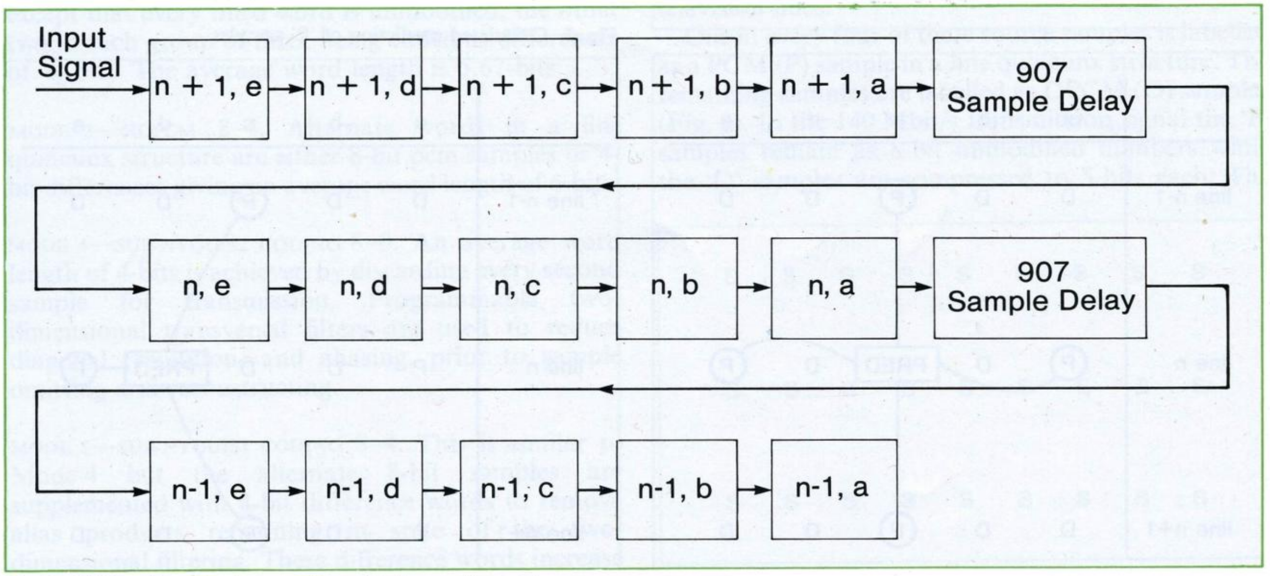


Fig. 5. Shift register implementation of 2-D sample matrix.

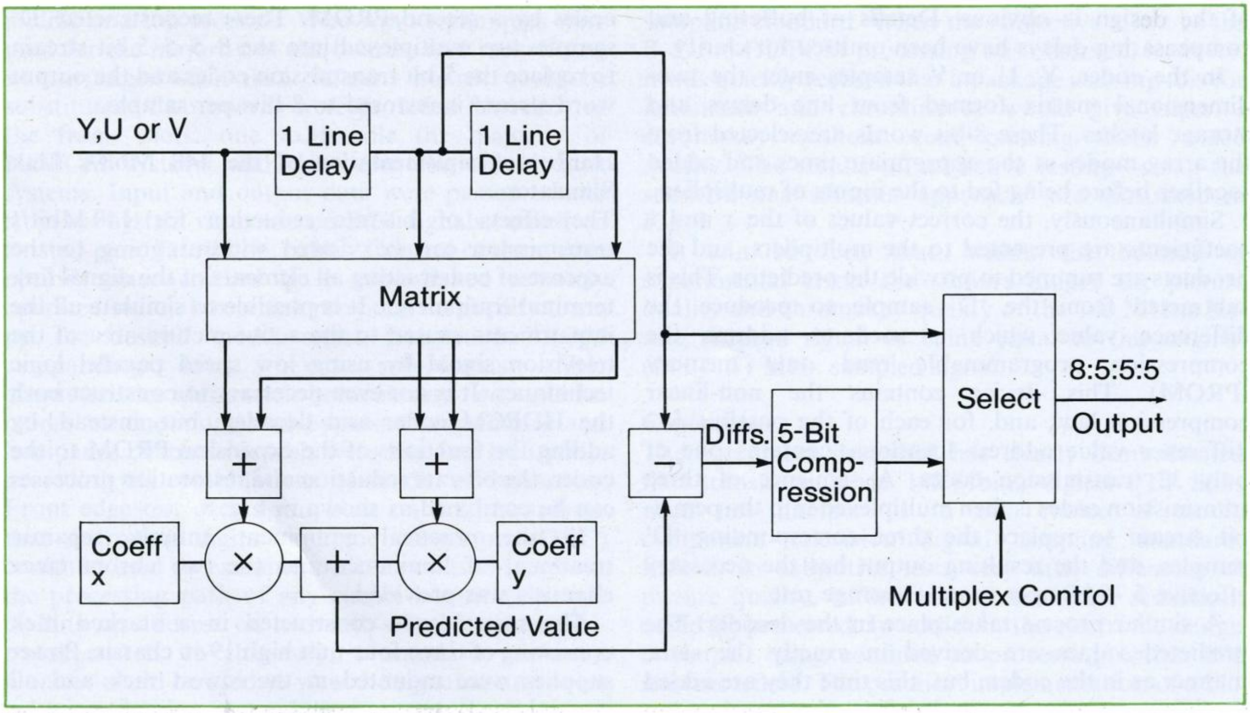


Fig. 6. Basic block diagram of 8 : 5 : 5 : 5 coder.

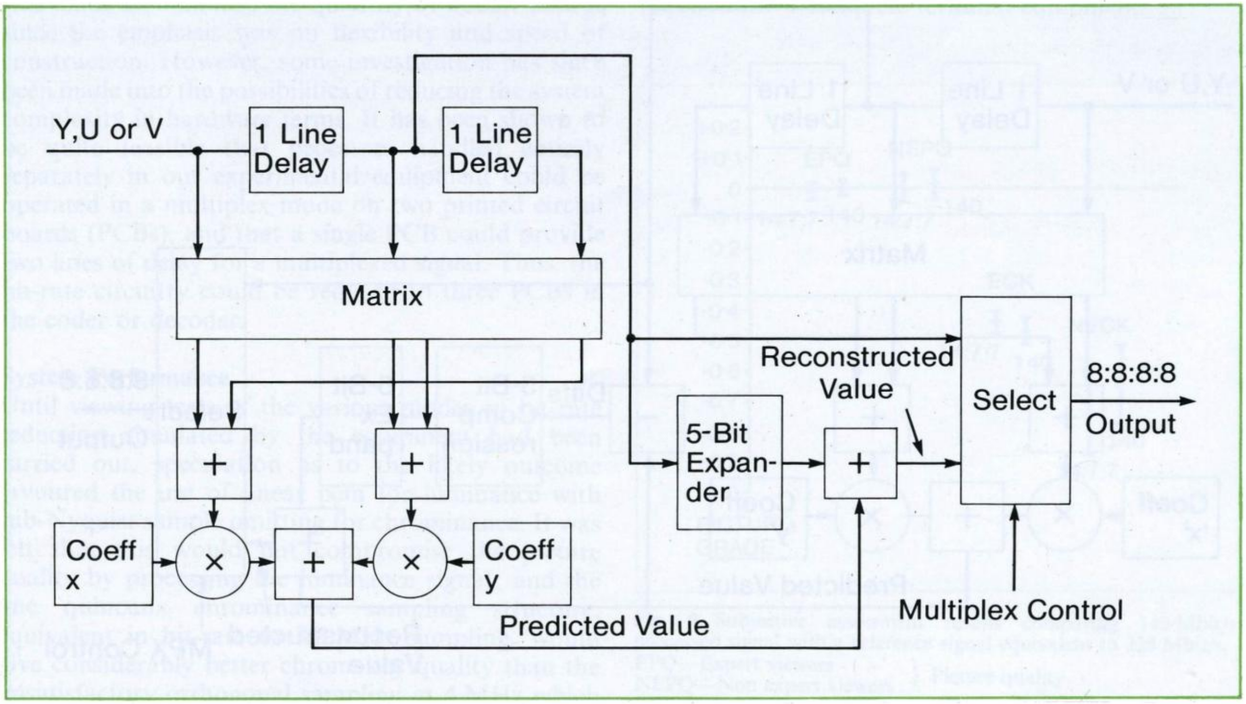


Fig. 7. Basic block diagram of 8 : 5 : 5 : 5 decoder.

of the design is obvious. Details of buffering and compensating delays have been omitted for clarity.

In the coder, Y, U or V samples enter the two-dimensional matrix formed from line delays and storage latches. These 8-bit words are selected from the array modes at the appropriate times and added together before being fed to the inputs of multipliers.

Simultaneously, the correct values of the x and y coefficients are presented to the multipliers, and the products are summed to provide the predictor. This is subtracted from the 'D' sample to produce the difference value which is used to address the compression programmable read only memory (PROM). This device contains the non-linear compression law; and, for each of the possible 512 difference value address locations, contains one of only 32 transmission codes. A sequence of three transmission codes is then multiplexed into the pcm 8-bit stream to replace the three corresponding 'D' samples, and the resulting output has the necessary effective 5.75-bits per sample average rate.

A similar process takes place in the decoder. The predicted values are derived in exactly the same manner as in the coder; but, this time they are added to difference values expanded from the transmission codes by a second PROM. These reconstructed 'D' samples are multiplexed into the 8–5–5–5 bit stream to replace the 5-bit transmission codes and the output word stream is restored to 8-bits per sample.

Hardware Implementation of the 140 Mbit/s Link Simulator

The effects of bit-rate reduction for 140 Mbit/s transmission can be viewed without going to the expense of constructing all elements of the digital link terminal equipments. It is possible to simulate all the impairments caused to the active picture area of the television signal by using low speed parallel logic techniques. It is not even necessary to construct both the HDPCM coder and decoder; but, instead, by adding the functions of the expansion PROM to the coder, the bit-rate reduction and restoration processes can be combined as shown in Fig. 8.

In the practical equipment entirely separate treatment of luminance and the two chrominance channels was provided.

The system was constructed in a stacked rack consisting of three four-unit high 19-in chassis. Power supplies were mounted in the lowest rack and all circuit boards were arranged in two rows of ten in the

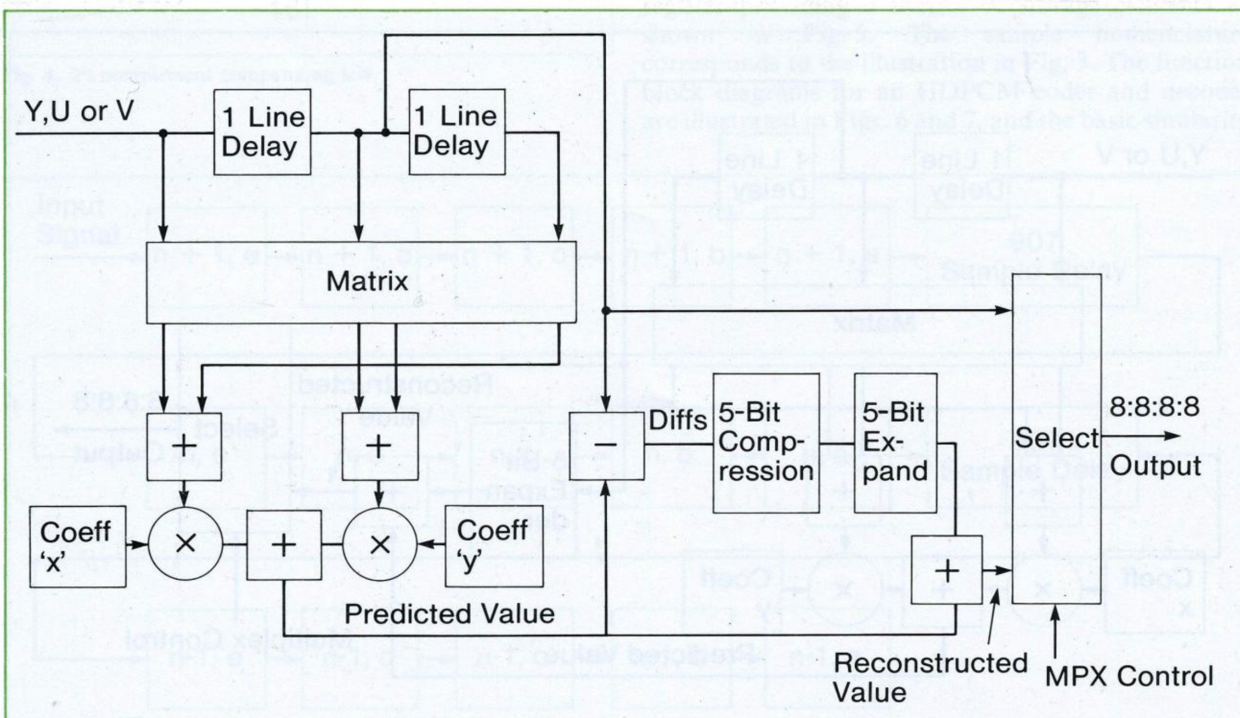


Fig. 8. Coder-decoder simulator.

middle and top sections. Wire-wrap techniques were used for the majority of the boards, and rack wiring was arranged such that different boards could be substituted. Two boards were provided for many of the frame slots, one to handle the majority of HDPCM systems and the other the sub-Nyquist systems. Input and output data were passed to and from the system via balanced ECL signals carried on twisted-pair, flat ribbon, cables. Control of the system options was achieved by a cable-linked remotely operated switch box which provided the facility to set up the desired mode, select alternative PROM companding laws, select sub-Nyquist filter characteristics and switch between any two systems, modes or one mode and bypass. Chrominance and luminance channel controls were handled separately, providing all combinations of modes of operation. Front edge test access was built onto all boards so that a 'roving digital/analogue convertor' (DAC) could be used to examine critical points throughout the processing path of any circuit board by simply dialling an address on a thumb-wheel hexadecimal-coded switch. In this way, a standard television monitor could be used to view the signal at any point in the processing, so speeding up development and fault finding procedures. In the hardware, no attempt was made to minimise the quantity of circuit boards since the emphasis was on flexibility and speed of construction. However, some investigation has since been made into the possibilities of reducing the system complexity in hardware terms. It has been shown to be quite feasible that processes handled entirely separately in our experimental equipment could be operated in a multiplex mode on two printed circuit boards (PCBs), and that a single PCB could provide two lines of delay for a multiplexed signal. Thus, the bit-rate circuitry could be reduced to three PCBs in the coder or decoder.

System Performance

Until viewing tests of the various modes of bit-rate reduction simulated by the equipment had been carried out, speculation as to the likely outcome favoured the use of linear pcm for luminance with sub-Nyquist sample omitting for chrominance. It was felt that this would not compromise the picture quality by processing the luminance signal; and the line quincunx chrominance sampling structure, equivalent in bit-rate to 3.5 MHz sampling, would give considerably better chromakey quality than the unsatisfactory orthogonal sampling at 4 MHz which was rejected in the move away from a (12 : 4 : 4) MHz sampling standard. When the subjective effect of the 8–5–5–5 HDPCM processing was observed, however, it was quickly realised that a package utilising this for luminance and chrominance would give superior chromakey without compromising basic video quality. The results of subjective testing—using the sensitive dual-stimulus approach[1] are illustrated in Fig. 9.

It can be seen that, within the bounds of experimental error, the viewers judged the picture quality and chromakey quality of the 140 Mbit/s bit-rate reduced signal as being equal to the original (14 : 7 : 7) MHz sampled signal.

Conclusion

The equipment described has clearly demonstrated that transmission via 140 Mbit/s digital PTT links need not hamper us in our choice of studio component television digital sampling standard. Indeed, in subjective tests very little difference of picture quality, before and after bit-rate reduction, could be perceived. In addition, the electronic design is simple, being confined to processing within one television field. The practical implementation of a bit-rate reduction coder or decoder would occupy three printed circuit boards of the total of fifteen to twenty required for a complete terminal equipment.

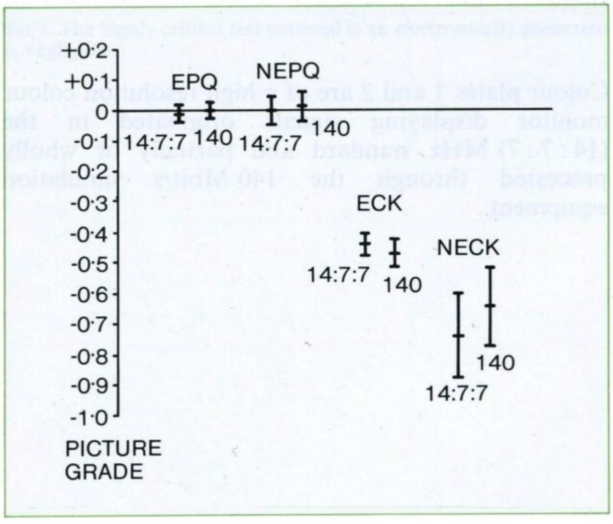


Fig. 9. Subjective assessment results comparing 140 Mbit/s processed signal with a reference signal equivalent to 228 Mbit/s.
EPQ—Expert viewers } Picture quality
NEPQ—Non expert viewers
ECK—Expert viewers } Chromakey quality
NECK—Non expert viewers

Plate 1. The left side of this scene is bit-rate reduced to 140 Mbit/s. The right side is at the equivalent of 228 Mbit/s.

Colour plates 1 and 2 are of a high resolution colour monitor displaying signals originated in the (14 : 7 : 7) MHz standard and partially or wholly processed through the 140 Mbit/s simulation equipment.

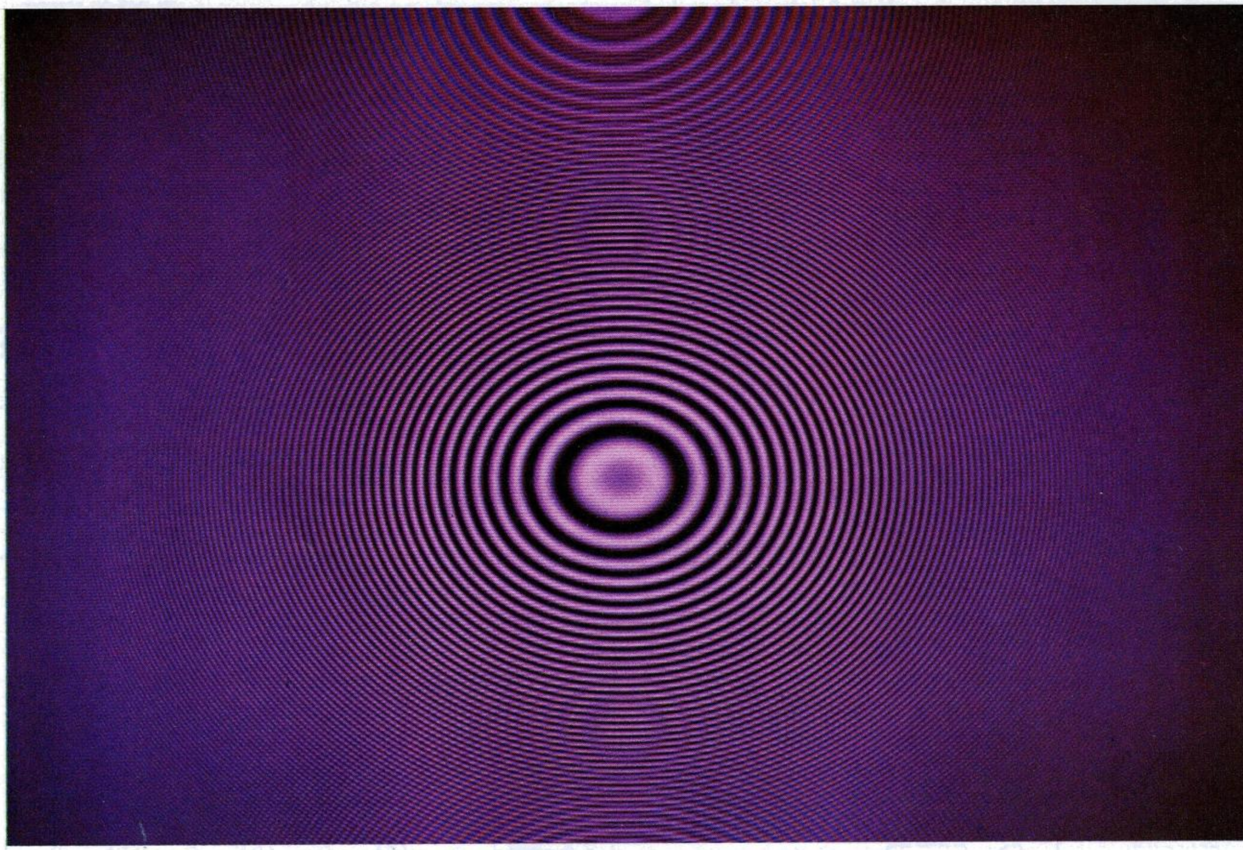


Plate 2. The result of bit-rate reduction from (14 : 7 : 7) MHz to 140 Mbit/s. The highly critical test material is an electronically generated magenta circular zone plate, and the effect of bit-rate reduction here is negligible.

APPENDIX A: OTHER BIT-RATE REDUCTION MODES

Mode 2—HDPCM 8–4½–4½

This works in a similar way to Mode 1 but with a different pattern 'P' and 'D' samples and with the predictor equations and weighting coefficients shown in Fig. 10 and Table 3.

$$\text{PRED}\,(n, b) = xP(n, a) + yP(n-1, b) + zP(n+1, c)$$

$$\text{PRED}\,(n, c) = xP(n, d) + yP(n+1, c) + zP(n-1, b)$$

This is a further two-dimensional HDPCM Mode similar to Mode 1 with alternate 'P' and 'D' samples. The predictor is more complex, however, utilising the weighted average of ten surrounding 'P' samples as shown in Fig. 11. Many sets of coefficients were programmed in the experimental equipment but the preferred one is listed in Table 4.

$$\text{PRED}\,(n, d) = w[P(n, a) + P(n, g)] + x[P(n-1, b) + P(n+1, b) + P(n-1, f) + P(n+1, f)] + y[P(n, c) + P(n, e)] + z[P(n-1, d) + P(n+1, d)]$$

The compression law used with this predictor has sixteen transmission codes carried by four bits.

TABLE 3: WEIGHTING COEFFICIENTS FOR 8–4½–4½

	X	Y	Z
Luminance	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{4}$
Chrominance	$\frac{3}{8}$	$\frac{5}{16}$	$\frac{5}{16}$

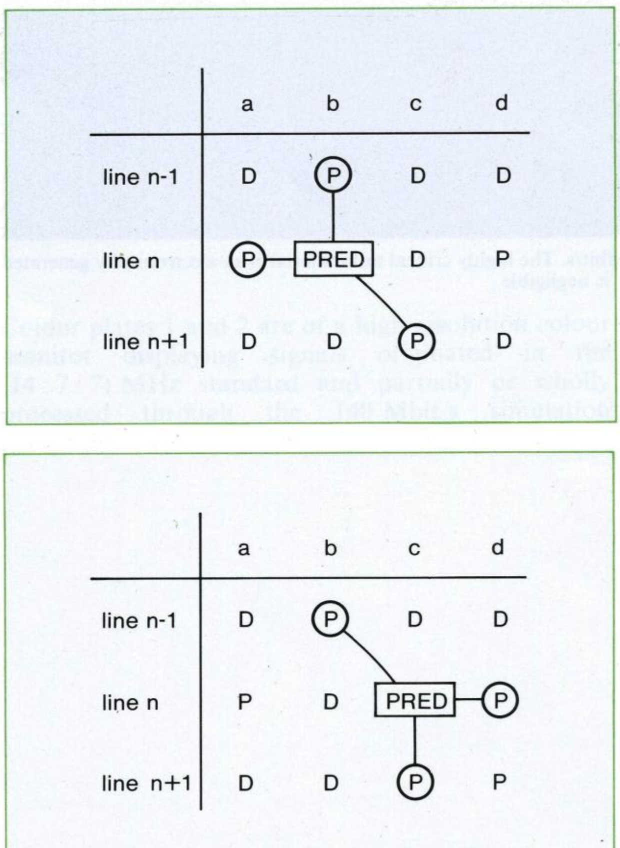


Fig. 10. HDPCM 8 : 4½ : 4½. 'P' and 'D' sample patterns.

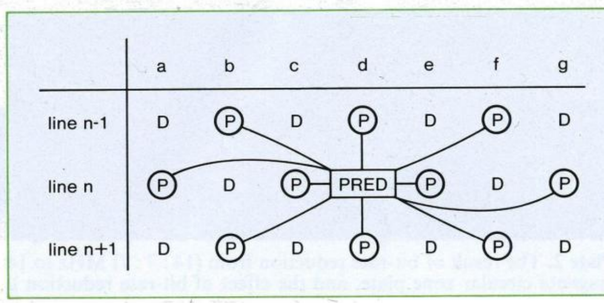


Fig. 11. Predictor for HDPCM 8 : 4.

TABLE 4: COEFFICIENTS FOR HDPCM 8–4

	W	X	Y	Z
Luminance & Chrominance	$\frac{1}{16}$	$-\frac{1}{8}$	$\frac{7}{16}$	$\frac{1}{4}$

The compression law used with this predictor has sixteen transmission codes carried by four bits.

APPENDIX B: COMPANDING LAWS

The difference values resulting from the 2's complement subtraction of the predictor from the appropriate 'D' sample have a possible range from −256 to +255, i.e. requiring nine bits to represent them in binary. In practice we use modulo 256 arithmetic where the number range is limited to 256 values. This means that each 8-bit binary number is used to carry two of the 512 possible difference values. The two values for a particular 8-bit code are simply its binary value and the binary value minus 256; for instance, 152 and −104. The ambiguity is easily resolved in the HDPCM decoder since only one of the pair of difference values can correctly reconstruct a 'D' sample within the allowed number range of the video signal. The implementation of modulo 256 arithmetic occurs automatically in our implementation when we ignore the carry out in the coder subtraction and decoder addition processes. The use of modulo 256 arithmetic means that the compression law must carry 256 different values in the same number of transmission codes in the same way as 512 difference values were carried in the 2's complement case. The improvement in picture quality is less than the 2 to 1 ratio would suggest, however, because the shape of the companding law must be modified to cope with the most critical of the pair of difference values carried by each 8-bit modulo 256 number. The typical shapes of the two types of compression law are illustrated in Fig. 12.

It can be seen that, at small difference values, both laws have a slope equivalent to 8-bit quantisation but the 2's complement law deviates from the linear nature more rapidly than does the modulo 256 law. The modulo 256 law has the peculiar property of also coding extreme difference values with 8-bit quantisation. In the worst case the modulo 256 law falls to about 4½-bit quantisation compared with the equivalent of 2½-bits for the 2's complement case. This reflects directly upon the errors made in reconstructing 'D' samples where, with modulo 256 arithmetic and law 'A', the maximum error is ±6 levels. With 2's complement arithmetic and law 'B' the error can be ±30 levels.

The modulo 256 law tabulated below was designed by using data obtained from distributions of difference values measured for typical picture material and some atypical test material (such as electronic zone plate).

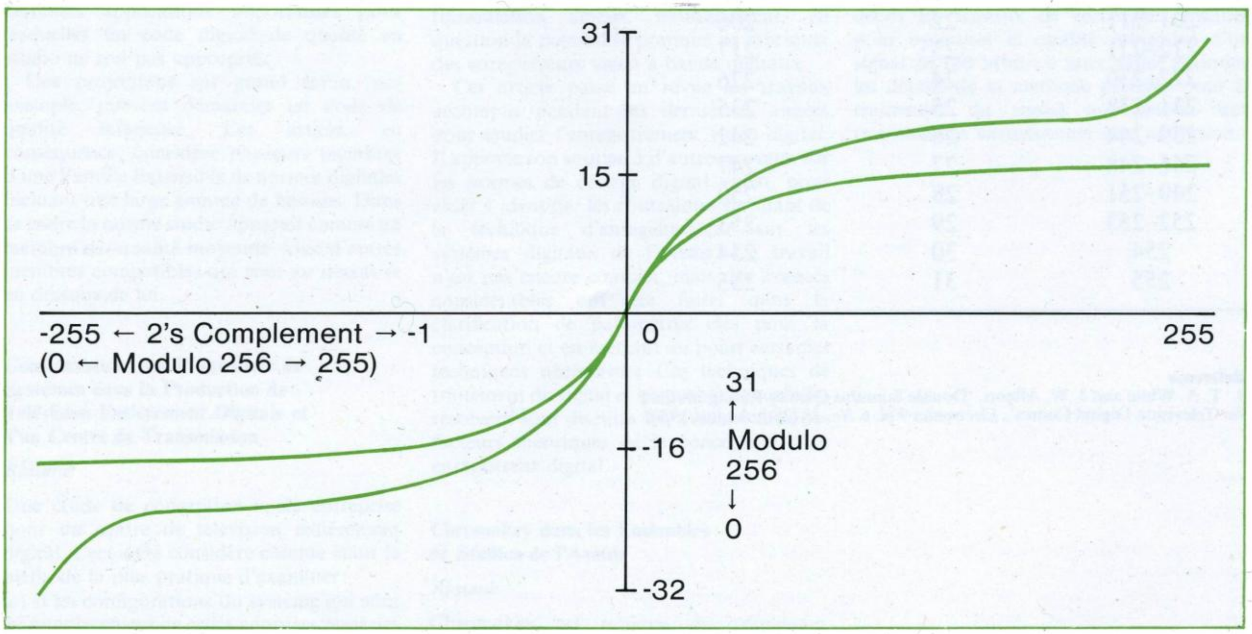


Fig. 12. Implementation of compression laws for 2's complement and modulo 256 arithmetic.

TABLE 5: MODULO 256 COMPRESSION/ EXPANSION LAW

INPUT DIFFERENCE VALUES	TRANSMISSION CODE	DIFFERENCE VALUE FOR RECONSTRUCTION
0	0	0
1	1	1
2–3	2	2
4–6	3	5
7–10	4	8
11–16	5	13
17–24	6	20
25–33	7	29
34–43	8	38
44–54	9	49
55–65	10	60
66–77	11	71
78–89	12	83
90–101	13	95
102–114	14	108
115–127	15	121
128–140	16	134
141–153	17	147
154–165	18	160
166–177	19	172
178–189	20	184
190–200	21	195
201–211	22	206
212–221	23	217
222–230	24	226
231–238	25	235
239–244	26	242
245–248	27	247
249–251	28	250
252–253	29	253
254	30	254
255	31	255

Reference

1. T. A. White and J. W. Allnott, 'Double Stimulus Quality Rating Method for Television Digital Codecs', *Electronics Vol. 6 No. 18* (28th August 1980).

L'Evolution vers des Systèmes Vidéo a Composants Codés

Résumé

Au cours des derniers mois des étapes importantes ont été franchies vers l'établissement d'un accord mondial sur une norme digitale pour les signaux de studios de télévision. Pour aboutir à un accord, et pour de solides raisons techniques et économiques, la nouvelle norme sera basée sur les signaux composants séparés qui constituent la base de divers signaux composites utilisés en ce moment. Cette communication résume l'histoire récente des discussions sur la normalisation digitale et considère les implications de la décision pour les signaux composants codés. Elle conclut que nous pouvons nous attendre à voir, dans les systèmes futurs, un changement progressif du codage des composites vers le codage des composants pour la distribution, et même pour la transmission de signaux de télévision.

Une Famille Extensible de Normes

Résumé

L'article précédent dans cette Revue Technique traite des facteurs ayant une influence sur le choix d'une norme mondiale d'échantillonage digital convenant aux composants de signaux de télévision pour une qualité en studio.

Toutefois, il est possible d'envisager certaines applications importantes pour lesquelles un code digital de qualité en studio ne soit pas approprié.

Des projections sur grand écran, par exemple, puevent demander un code de qualité inférieure. Cet article, en conséquence, considère plusieurs membres d'une Famille Extensible de normes digitales incluant une large gamme de besoins. Dans ce cadre la norme studio apparaît comme un membre de 'qualité moyenne' avec d'autres membres compatibles qui sont au dessus et en dessous de lui.

Considerations sur l'Ingénierie de Systèmes dans la Production de Télévision Entièrement Digitale et d'un Centre de Transmission

Résumé

Une étude de conception a été entreprise pour un centre de télévision entièrement digital. Ceci a été considéré comme étant la méthode la plus pratique d'examiner:
(*a*) si les configurations du système qui sont en remplacement de celles adoptées pour des environements analogiques seraient avantageuses,
(*b*) d'évaluer ce que seraient les conséquences de ces configurations de remplacement sur les caractéristiques du nouveau matériel digital,
(*c*) de déterminer comment la phase d'introduction du système digital dans le fonctionnement de la télévision pourrait être effectué au mieux, et
(*d*) d'évaluer l'aptitude d'un jeu particulier de paramètres de codage digital à leur utilisation dans le contexte d'un système général.

Cette communication met l'emphase sur ces aspects de l'étude qui concernent la configuration du système et les caractéristiques souhaitables en résultant du matériel digital, spécialement en ce qui concerne l'adoption des normes pour le multiplexage pour le vidéo, l'audio et les impulsions à l'intérieur de l'installation.

Enregistreurs Vidéo sur Bande Digitaux vis-a-vis des Signaux à Composants Codes

Résumé

L'enregistreur vidéo sur bande est devenu un élément vital dans la production moderne de programmes de télévision. Il s'est développé pour remplir divers rôles reflétant les nombreux aspects du procédé de production. L'éventualité que la technique digitale puisse, peut-être dans le proche avenir, commencer à dominer la transmission amène, naturellement, en question la possibilité pratique de fabriquer des enregistreurs vidéo à bande digitaux.

Cet article passe en revue les travaux accomplis pendant les dernièrnes années pour étudier l'enregistrement vidéo digital. Il apporte son soutien à d'autres travaux sur les normes de codage digital vidéo, pour aider à identifier les contraintes résultant de la technique d'enregistrement sur les systèmes digitaux de l'avenir. Ce travail n'est pas encore complet, mais des avances considérables ont été faites dans la clarification de paramètres clés pour la conception et en mettant au point certaines techniques nécessaires. Ces techniques de traitement du signal et les problèmes qu'elles résolvent sont discutés ici, avec certains des facteurs théoriques de la conception d'un enregistreur digital.

Chromakey dans les Ensembles de Studios de l'Avenir

Résumé

Chromakey est typique du traitement complexe du signal nécessaire dans les studios modernes de télévision, et son utilisation comme technique de production est d'une importance grandissante. Quantités de discussions sur les normes d'échantillonnage digital en studio se sont concentrées sur les besoins de largeur de bande pour le canal couleur. L'effet que ces discussions auront sur la performance de la chromakey digitale est discuté en cet article. Deux essais expérimentaux sont discutés, à savoir, la conception d'un matériel de chromakey digital et une étude des relations entre la norme d'échantillonage et la performance de chromakey.

Réduction du Taux de Bits pour les Liaisons 140 Mbit/s

Résumé

Les progrès récents de la technique, spécialement dans le domaine des câbles en fibres optiques ont amené au premier plan l'introduction planifiée de liaison digitales de grande capacité.

La liaison 140 Mbit/s au niveau du multiplexage au quatrième ordre de la norme Européenne peut permettre la conversion économique de tout le réseau en fonctionnement digital—peut-être dans une décenie. Pour les émetteurs, un réseau entièrement digital de liaisons à 140 Mbit/s nous permettrait de contempler la qualité supérieure possible au moyen de la transmission de signaux de télévision digitaux en forme de composants. Le texte décrit les travaux de recherche effectués pour optimiser la qualité subjective d'un signal de 140 Mbit/s à taux réduit et donne les détails de la méthode préférée pour le traitement du signal qui permet une transmission virtuellement sans altération.

Die Entwicklung von Video-Systemen mit Signalkomponentenkodierung

Übersicht

In den letzten Monaten wurde wichtige Schritte zur weltweiten Einigung über eine Norm für digitale Fernsehstudiosignale unternommen. Zur Erzielung einer Einigung und aus fundierten technischen und ökonomischen Gründen wird die neue Norm auf den getrennten Signalkomponenten basieren, die heute verwendet werden. In dieser Abhandlung wird der neuere Stand der Diskussion zum Thema der Digital-Normierung zusammengefaßt. Des weiteren werden die wahrscheinlichen Folgen der Tendenz zur Verwendung der Signalkomponentenkodierung betrachet. Es eribt sich, daß wir bei zukünftigen Systemen einen progressiven Übergang von Verbundkodierung zu Komponentenkodierung erwarten können, und zwar nicht nur bei der Distribution sondern auch bei der Übertragung von Fernsehsignalen.

Eine erweiterungsfähige Normenreihe

Übersicht

Im vorhergehenden Artikel in dieser Ausgabe des Technical Review werden Faktoren betrachtet, die die Wahl einer geeigneten, weltweit gültigen Norm für die Digitalabtastung von zusammengesetzten Studio-Fernsehsignalen beeinflussen. Es lassen sich jedoch einige wichtige Anwendungsbereiche denken, bei denen eine Digitalkodierung von Studioqualität nicht angebracht ist.

Großbildschirm-Sichtgeräte könnten beispielsweise einer Kode von geringerer Qualität benötigen. Aus diesem Grunde betrachtet dieser Artikel mehrere Mitglieder einer möglichen, erweiterungsfähigen Reihe von Digitalnormen, die mehrere Anwendungsgebiete erfassen. In diesem Rahmen erscheint die Studionorm als die Norm für 'mittlere' Qualität, während entsprechende weniger anspruchsvolle oder anspruchsvollere Norman unter- bzw. übergeordnet sind.

Systemtechnische Betrachtungen für das Volldigital-Fernsehproduktions- und Sendezentrum

Übersicht

Es wurde eine Konstruktionsstudie für ein volldigitales Fernsehzentrum durchgeführt. Eine solche wurde als geeeignetstes Verfahren zur Untersuchung der nachfolgenden Fragen erachtet:

(*a*) Sind die Systemkonfigurationen, die als Alternativen für Analogverhältnisse eingesetzt werden, von Vorteil?
(*b*) Was für Folgen haben solche Alternativkonfigurationen für die Spezifikationen der neuen Digitalausrüstungen.
(*c*) Wie laßt sich die Einführung des Digitalsystems in den Fernsehbetrieb am besten erreichen, und
(*d*) Wie geeignet ist ein gegebener Parametersatz für die Digitalverschlüsselung im Rahmen eines Gesamtsystemkonzepts.

Diese Abhandlung betont diejenigen Gesichtspunkte der Untersuchung, die sich auf die Systemkonfiguration und die sich daraus ergebenden, wünschenswerten Merkmale der Digitalausrüstung beziehen, besonders was ihre Bedeutung für die Anerkennung von Normen für die Mehrfachübertragung des Videosignals, des Audiosignals und der Impulse in der Installation betrifft.

Digital-Videorecorder für Signale mit Komponentenkodierung

Übersicht

Aus der moderne Fernsehprogrammproduktion ist heute der Videorecorder nicht mehr wegzudenken. Er ist heute so weit entwickelt, daß er eine Reihe von Rollen erfüllt, die die unterschiedlichen Gesichtspunkte der Fernsehproduktion berücksichtigen. Die Tatsache, daß die Digitaltechnologie in der nahen Zukunft anfangen kann, das Rundfunk- und Fernsehwesen zu dominieren, führt natürlich zwangsläufig zu der Frage der Möglichkeit der Digitalen Fernsehaufzeichnung.

Diese Abhandlung vermittelt eine Übersicht über die Arbeit der letzten Jahre zur Untersuchung der Möglichkeiten der digitalen Fernsehaufzeichnung. Diese Arbeit ergänzt gleichzeitig andere Arbeiten auf dem Gebiet der Digital-Videoverschlüsselung und deren Normierung, die die Identifikation von Beschränkungen für zukünftige Digital-Studiosysteme aufgrund der Aufzeichnungstechnologie erleichtern sollen. Diese Arbeiten sind noch nicht abgeschlossen, jedoch wurden bei der Klärung wesentlicher Konstruktionsparameter und bei der Entwicklung bestimmter Arbeitsverfahren wesentliche Fortschritte erzielt. Diese Signalverarbeitungsverfahren und die durch sie gelösten Probleme werden hier zusammen mit einigen der theoretischen Faktoren der Digital-Aufzeichnungsgeräte erörtert.

Chromakey Verfahren in zukünftigen Studiosystemen

Übersicht

Das Chromakey-Verfahren ist typisch für die komplizierten Signalverarbeitungsverfahren, die heute in modernen Fernsehstudios erforderlich sind, und der Einsatz dieses Verfahrens als Produktionsverfahren gewinnt laufend an Bedeutung. Die Diskussion der Digital-Studionormen konzentrierte sich hauptsächlich auf die Anforderungen in bezug auf die Farbkanal-Bandbreite. Der Einfluß dieser auf die Leistung des Digital-Chromakey-Verfahrens wird hier erörtert. Es werden zwei Versuchsreihen beschrieben, nämlich die Konstruktion einer Digital-Chromakey-Anlage und die Untersuchung der Beziehung zwischen Abtastleistung und Chromakey-Leistung.

Reduktion der Bitgeschwindigkeit bei Verbindungen mit 140 Mbit/s

Übersicht

Neue Fortschritte auf dem Gebiet der Technologie, insbesondere auf dem Gebiet der Lichtleiter, haben die geplante Einführung von Digitalverbindungen mit hoher Kapazität beschleunigt.

Die 140 Mbit/s-Verbindung mit der Mehrfachübertragung vierter Ordnung nach der europäischen Norm könnte die wirtschaftliche *Umrüstung* des gesamten Digitalnetzes innerhalb von zehn Jahren ermöglichen. In der Rundfunk- und Fernsehtechnik würde uns ein Volldigitalnetz aus Verbindungen mit 140 Mbit/s in die Lage versetzen, die Möglichkeit der Nutzung der überlegenen Qualität bei Übertragung von Fernsehsignalen in Digitalform in Betracht zu ziehen. Diese Abhandlung beschreibt die Untersuchungsarbeiten, die zur Optimierung der subjektiven Qualität eines reduzierten Signals von 140 Mbit/s durchgeführt wurden und vermittelt Einzelheiten über das bevorzugte Verfahren der Signalverarbeitung, das praktisch störungsfreie Übertragung über die Verbindungen ermöglicht.

La Evolución Hacia Sistemas Video de Componentes Codificadas

Resumen

En los últimos meses se han tomado medidas importantes para establecer un acuerdo de ámbito mundial sobre un estándar digital para señales de estudio de televisión, Para conseguir el acuerdo, y por obvias razones técnicas y económicas, el nuevo estándar estará basado en las señales componentes separadas que forman la base de las diversas señales compuestas que se usan actualmente. Este artículo compendia la historia reciente de las discusiones sobre normalización digital, y considera las consecuencias del cambio hacia las señales de componentes codificadas. La conclusión es que podemos esperar, en los sistemas futuros, un cambio progresivo de la codificación compuesta a la de componentes para la distribución, e incluso para la transmisión, de señales de televisión.

Una Familia de Normas Extensible

Resumen

El artículo anterior de esta Revista Técnica trata de los factores que influyen en la elección de una norma de muestreo digital mundial adecuada para señales de televisión de componentes de calidad de estudio. Sin embargo, se pueden prever algunas aplicaciones importantes para las que un código digital de calidad de estudio no es apropiado.

Las presentaciones de pantalla grande, por ejemplo, pueden requerir un código de calidad inferior. Por tanto, este artículo considera varios miembros de una posible familia extensible de normas, englobando una amplia gama de requisitos. Dentro de este marco, la norma de estudio aparece como un miembro de 'calidad media' con otros miembros compatibles por encima y por debajo.

Consideraciones de Ingeniería de Sistemas en el Centro de Transmisión y Producción de Televisión Completamente Digital

Resumen

Se ha emprendido un estudio de diseño para un centro de televisión completamente digital. Este fue el método más práctico de exploración:

(*a*) viendo si las configuraciones de sistemas alternativos a los adoptados para entornos analógicos serían ventajosas;

(*b*) evaluando cuales serían las consecuencias de estas configuraciones alternativas en las especificaciones de equipo digital nuevo;

(*c*) determinando cómo se podría conseguir mejor la transición de la introducción del sistema digital en una operación de televisión, y

(*d*) valorando la conveniencia de un conjunto de parámetros de codificación digital para el uso en el contexto de un sistema general.

Este artículo destaca aquellos aspectos del proyecto correspondientes a la configuración del sistema y a las características resultantes deseables del equipo digital, especialmente en lo que se refiere a la adopción de normas para multiplexar el video, el audio y los impulsos dentro de la instalación.

Videógrafos Digitales para Señales de Componentes Codificadas

Resumen

El videógrafo se ha convertido en un elemento vital en la producción moderna de programas de televisión. Ha evolucionado para desempeñar varios papeles que reflejan numerosos aspectos del proceso de producción. La posibilidad de que la tecnología digital pueda, tal vez dentro de un futuro próximo, empezar a dominar la radiodifusión, naturalmente trae a colación la factibilidad práctica de los videógrafos.

Este artículo revisa el trabajo efectuado durante los últimos años para investigar la videograbación. El mismo apoya otro trabajo sobre normas de codificación de video digital, para ayudar a identificar las limitaciones, impuestas por la tecnología de grabación, en los sistemas de estudios digitales futuros. Aunque no se ha terminado aún este trabajo, se ha efectuado un avance considerable para clasificar parámetros clave de diseño y para desarrollar ciertas técnicas necesarias. Aquí se discuten estas técnicas de proceso de señales y los problemas que resuelven, junto con algunos de los factores teóricos del diseño de videograbadores digitales.

Cromaticidad (Chromakey) en Sistemas de Estudio Futuros

Resumen

La cromaticidad es típica del complejo proceso de señales necesario en los estudios de televisión modernos, y su empleo como técnica de producción es de importancia creciente. Gran parte de la discusión sobre normas de muestreo de estudio digitales se ha centrado en los requerimientos de ancho de banda del canal de color. Se discute aquí el efecto que tendrán las mismas en el rendimiento de la cromaticidad digital. Se describe dos investigaciones experimentales, que son: el diseño de un equipo de cromaticidad digital y un estudio de la relación entre las normas de muestreo y el rendimiento de la cromaticidad.

Reducción de la Tasa de Bitios para Enlaces de 140 Mbit/s

Resumen

Recientes avances tecnológicos, particularmente en el área de los cables de fibra óptica, han impulsado la introducción planificada de enlaces digitales de alta capacidad.

El enlace de 140 Mbit/s a nivel de multiplex de cuarto orden de estándar europeo, puede permitir la conversión económica de la red total a operación digital, tal vez dentro de una década. Para los radiodifusores, una red digital completa de enlaces de 140 Mbit/s permitiría contemplar la calidad superior posible por transmisión de señales de televisión digitales en forma de componente. El texto describe trabajos de investigaciones efectuadas para optimizar la calidad subjectiva de una señal de tasa reducida de 140 Mbit/s y da detalles del método preferido de proceso de señales que permite virtualmente una transmisión de enlace impecable.

IBA TECHNICAL REVIEW

1 Measurement and Control*

2 Technical Reference Book (3rd Edition)*

3 Digital Television*

4 Television Transmitting Stations*

5 Independent Local Radio*

6 Transmitter Operation and Maintenance*

7 Service Planning and Propagation*

8 Digital Video Processing—DICE*

9 Digital Television Developments*

10 A Broadcasting Engineer's Vade Mecum

11 Satellites for Broadcasting

12 Techniques for Digital Television

13 Standards for Television and Local Radio Stations*

14 Latest Developments in Sound Broadcasting*

15 Microelectronics in Broadcasting Engineering

*Out of Print

IBA

INDEPENDENT
BROADCASTING
AUTHORITY

Printed by The Whitefriars Press Ltd, London and Tonbridge

IBA INDEPENDENT BROADCASTING AUTHORITY